Three Essays in Development Economics

Kyle McNabb
Submitted for the degree of Doctor of Philosophy

Heriot-Watt University
School of Management and Languages
September 2016

Abstract

This thesis contains three empirical studies on different facets of development economics. Two of the chapters specifically focus on Benin, a country that has not often been studied in the development literature. The first of these studies is an investigation into the sustainability of Beninese Rotating Savings and Credit Associations. Work such as Besley *et al.* (1993) or Anderson *et al.* (2009) has questioned how such groups can overcome the incentives for individuals to default, theorising that the institutional design of the groups can play an important role. Using household survey data collected in 2004 and 2006, chapter 2 presents a first test into many of the theories outlined in the theoretical literature. The second study examines primary school attendance rates in Benin. Despite almost unparalleled increases in attendance rates since 1990, the country has remained virtually ignored in the literature. The study, found in chapter 3, attempts to improve on existing related studies by employing a multilevel model in order to account for higher level variance in the data. The results help to shed light on the stark regional disparities in attendance that continue to exist across Benin's 77 communes. Finally, chapter 4 uses the recently released ICTD-UNU WIDER Government Revenue Dataset in order to revisit some recent results on the relationship between tax structures and economic growth. Recent work in this field has offered little or no evidence for developing countries, yet it is in such countries where the greatest changes in tax structure have not only been seen over the past 30 years but will likely be seen in coming years. The study examines the impact of revenue neutral changes in tax structure on per capita GDP growth rates. Results suggest that previous findings are not applicable to countries at all levels of development and as such urge caution with regard to some recent policy advice that is based on previous findings.

DECLARATION STATEMENT

(Research Thesis Submission Form should be placed here)

Table of Contents

List of Figures

List of Tables

Glossary

ADF	Augmented Dickey-Fuller
AME	Average Marginal Effect
CFA	West - African Franc
CIT	Corporate Income Tax
CNE	*Caisse Nationale d'Epargne*
DFE	Dynamic Fixed Effects
DHS	Demographic and Health Survey
ECM	Error Correction Model
EU	European Union
FINCA	Foundation for International Community Assistance
GDP	Gross Domestic Product
GER	Gross Enrolment Rate
GFS	Government Finance Statistics
GPI	Gender Parity Index
GRD	Government Revenue Dataset
HLM	Hierarchical Linear Model
HLP	High Level Panel
ICC	Intra-Class Correlation Coefficient
ICTD	International Centre for Tax and Development
ILO	International Labour Organisation

IMF	International Monetary Fund
INSAE	*Institute National de la Statistiqe et de l'Analyse Economique*
LCU	Local Currency Units
MDG	Millennium Development Goal
MEM	Marginal Effect at Mean
MFI	Microfinance Institution
MG	Mean Group
MLE	Maximum Likelihood Estimation
MLM	Multilevel Linear Model
MQL	Marginal Quasi-Likelihood
NAFTA	North American Free Trade Agreement
NER	Net Enrolment Rate
NGO	Non-Governmental Organisation
ODA	Official Development Assistance
ODI	Overseas Development Institute
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PIT	Personal Income Tax
PMG	Pooled Mean Group
PSU	Primary Sampling Unit
PTR	Pupil Teacher Ratio

PQL	Penalised Quasi-Likelihood
RN	Revenue - Neutral
ROSCA	Rotating Savings and Credit Association
SDG	Sustainable Development Goal
SSA	Sub-Saharan Africa
SURE	Seemingly Unrelated Regression Equation
UIS	UNESCO Institute for Statistics
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNU-WIDER	United Nations University World Institute for Development Economics Research
USD	United States Dollar
VAT	Value-Added Tax
VPC	Variance Partition Coefficient
WDI	World Development Indicators
WEO	World Economic Outlook

Chapter 1. Introduction

This thesis contains three empirical studies, each of which attempts to answer questions pertinent to key development issues today. Whilst the topics considered are somewhat distinct from one another, commonality can be found in the use of new, or more appropriate data, to shed light on existing questions or theories.

Considering the importance of data, there has recently been an increase in questions over the quality of statistics produced by governments, and in turn used by researchers, that have often been used to inform policy conclusions in developing countries (e.g. Jerven, 2013; Devarajan, 2013). Good quality data is of central importance to development economists; without it we might not know where we stand, where to aim, or how to get there. Concurrently, the High Level Panel (HLP) of the United Nations (UN) has recently called for a 'data revolution' in order to improve the quality and accessibility of key metrics that measure development progress (UN HLP, 2013). One need only look at the lengthy list of Sustainable Development Goal (SDG) indicators as an illustration of the need for good quality data on a wide variety of metrics. Indeed, over 200 indicators are proposed for the 17 goals.

The three studies contained herein each take different approaches to ensure that the best data, or the best use of data, is employed in order to arrive at the reliable empirical conclusions. Chapter 2 relies on first-hand household survey data, collected according to a random process in Benin in 2004 and 2006; the advantages of random sampling are well understood and should lead to a representative sample being drawn from the population. The next chapter combines household survey data from the Beninese Demographic and Health Surveys (DHS) with administrative data from the Beninese institute for statistics (INSAE); the DHS is widely used, understood as representative and its methods are transparent and trusted amongst the research community. With regard to the key metric measured – primary school enrolment – the DHS data paints a relatively different picture to official statistics from, for example UIS (UNESCO Institute for Statistics). The chapter includes a discussion of why this might be the case and how to best make sense of the disparity. Finally, the fourth chapter makes use of the recently released *Government Revenue Dataset* (GRD), developed at the International Centre for Tax and Development (ICTD) and UNU-WIDER. This dataset is the result of work to synthesise official revenue statistics from a number of sources such as the IMF, World Bank and OECD. As in chapter

3, a discussion highlights the disparities across different sources and seeks to explain how the GRD overcomes these to ensure more analytically accurate data is presented.

With regard to the context of the thesis, two of the three studies present evidence from Benin, a country that has been studied in comparatively lesser detail than many of its African neighbours, but provides an interesting case study in development. A small Francophone West African country of around 11 million inhabitants, Benin is bordered by Togo, Burkina Faso, Niger and Nigeria. Benin has seen rapid development since the fall of the communist regime in the late 1980's. Its economy is dominated by the service and agricultural sectors, with cotton representing the largest export. The country is, today, ruled by a comparatively stable democracy by African standards, and has seen stable growth levels in recent years. Yet, Benin is still faced with many challenges: poverty remains high and the gains from development have not been evenly shared geographically. Chapter 3 of this thesis studies an example of such uneven development in depth; whilst national primary school enrolment rates have seen almost unparalleled improvements since 1990, many regions have not shared in this progress and still lag behind. Furthermore, national completion rates have fallen somewhat, reflecting some of the difficulties that accompany such rapid progress. The country has also experienced rapid population growth since 1990, with the population having more than doubled in just 20 years. This has meant that even stable high growth rates of over 5% in recent years have not been sufficient to reduce poverty levels; the most recently available data suggests that over one third of Benin's citizens still live in poverty – the *poverty headcount ratio* in 2011 was 36.2% (WDI, 2016). Furthermore, life expectancy remains low and child mortality high, with around 100 deaths per 1000 births, as of 2015.

1.1 Outline of Chapters & Summary of Thesis

The first study of this thesis is entitled *Enforcement Problems in ROSCAs: An Empirical Investigation* and is found in chapter 2. This work examines the sustainability of an informal savings group, the Rotating Savings and Credit Association (ROSCA), in Benin. Despite the marked growth in popularity of microfinance or mobile banking, many of the world's poor still have no choice but to to save with friends, family or colleagues in informal arrangements such as ROSCAs. This is particularly relevant in Benin where, despite entrenched cultural habits of saving, the vast majority of savings in the country were held in ROSCAs or with itinerant bankers (Helms *et al.*, 2005). Indeed, in 2004, just 10% of the country's deposits were held within the formal banking system. Often individuals use ROSCAs in order to commit to a savings habit or to finance the purchase

of a durable good. However, if the group itself should experience default, then this will cause a negative income shock for those members who potentially lose their savings. Much has been written of the remarkable stability of ROSCAs (*inter alios:* Besley *et al*., 1993; Anderson *et al*. 2009) in the face of individual incentives to default, yet the theories outlined in the literature that describe this have, until now, gone untested. Using a unique panel dataset collected in urban Benin in 2004 and 2006, a group-level analysis tests many of the theories outlined in relation to how ROSCAs can be designed in order to influence sustainability of the arrangement and as such avoid default. The results highlight that Beninese ROSCAs are inherently unstable; in the two years between the surveys, one in three of the groups surveyed had experienced a case of non-payment or outright default. The analysis suggests that institutional features such as the order in which the collective savings is allocated, the ruling structure, the frequency of meetings and the sociocultural makeup of the group are all related to the likelihood of a group experiencing enforcement problems.

The second study is entitled *Exploring Regional and Gender disparities in Beninese Primary School Attendance: A multilevel approach* and can be found in chapter 3. Benin provides an interesting case study with regard to education and development, but has not yet been considered in the literature surveying school enrolment. By 1990, an economic crisis, coupled with the former socialist regime's failed attempts at reform had left the country with some of the worst primary school enrolment rates in the world: Gross enrolment was less than 50% and only one in three girls were attending. Since then, however, Benin has seen almost unparalleled progress. Over the next 20 years, gross and net enrolment ratios soared and the gender gap was all but eliminated. However, this progress has not occurred across the board. Regional disparities still persist today: the latest round of the Beninese DHS highlights that net attendance rates in many of Benin's 77 communes were close to 90%, but in others, only 20-30% of primary school age children were attending. Furthermore, significant differences in primary school attendance still exist with regard to gender. For the school year considered, 2005-06, girls were still around 9% less likely to attend than boys. Whilst a number of recent studies in the field have acknowledged that factors at the household or community level might play a role in predicting the likelihood that a child is sent to school, very few explicitly account for this in the econometric framework. Combining DHS data with detailed administrative data from INSAE, the empirical part uses a multilevel model in order to account for clustering at the household and commune level. The results highlight the importance of such higher level

factors. In particular, I find that as the distance to school increases, the likelihood that boys who worked in the field attended school was lower than for boys who did not work. This particular result highlights that in a context where direct costs and benefits of sending children to school might not be easily observed or estimated, the opportunity cost still appears to be taken into consideration. A random intercepts model highlights that after accounting for individual, household and community factors, much of the unexplained variance in primary school attendance lies at the household level. A random slopes model then pinpoints those communes where increases in household income, or reductions in the cost of schooling, might realise the largest future gains in primary school attendance.

The third study, *Tax Structures Economic Growth and Development* found in chapter 4, makes use of the recently developed ICTD UNU-WIDER Government Revenue Dataset (GRD) in order to revisit questions surrounding the effects of tax structure on economic growth. This dataset overcomes a number of issues with data from existing sources in order to achieve notable gains in analytical accuracy and a more complete series for many countries. The study represents a challenge to recent work in the field that suggests that consumption and property taxes are more growth-friendly than are income taxes (*inter alios:* Arnold *et al*., 2011, Acosta-Ormaechea and Yoo, 2012). Whilst the results of Pooled Mean Group regressions suggest that this is the case for high income countries, this is not generalisable; different taxes appear to have heterogeneous effects on economic growth at different income levels. This finding in itself is important: recent output from the IMF (IMF, 2011; IMF 2015) promotes the view that personal or corporate income taxes are more harmful for GDP growth than consumption taxes. However, crucially, these results are based on studies that only include OECD, high and some middle - income countries. If policy advice is to be provided to developing countries on the effects of tax policy on growth, then the knowledge base underpinning this advice should at least be informed by results from low or middle income countries. The work presented in chapter 4 represents a first step providing evidence in that respect. The study is also the first to explicitly consider the effects on growth rates of replacing trade with domestic consumption taxes, something that has become the standard IMF 'policy prescription' (IMF 2011) for developing countries. Results suggest that there have not always been positive effects of such trade liberalisation on economic growth.

Chapter 5 concludes by providing further discussion of the potential policy implications, or avenues for further research that might emerge as a result of the results presented herein.

Chapter 2. Enforcement Problems in ROSCAs: An Empirical Investigation

2.1 Introduction

Informal financial institutions such as rotating savings and credit associations (ROSCAs) are widespread in the developing world, particularly where individuals have limited access to formal credit or savings markets.[1] Despite the proliferation of alternatives such as microcredit or mobile banking, many people still choose to keep some or all of their savings in a ROSCA. For example, when Dupas *et al.* (2012) offered to remove the cost of opening a bank account for rural Kenyans, only 18% actively used the accounts whilst over 40% of their sample continued to use a ROSCA. Yet the sustainability of ROSCAs has often puzzled economists; the incentives for the self-seeking individual to renege on payments either before or after she has received the collective savings often outweigh those to remain loyal to the arrangement. There are other disadvantages of saving in a ROSCA compared to some alternative means, such as saving in autarky; they require an individual to commit to an inflexible savings pattern, they do not provide interest and cannot be legally enforced. Numerous studies have shed light on the importance of social connectedness between members, who front a kind of social collateral (Besley *et al.*, 1993) to their fellow members as security. The threat of social sanctions (in practice: obtaining a reputation for being unreliable, or exclusion from the group and future groups in the region) brings with it a large cost to default and thus an incentive to remain loyal to one's peers. At the same time, many ROSCAs also serve a social purpose. Members may offer food or drinks to each other during meetings, celebrate together on special occasions or share contacts and potential business opportunities. Such occurrences not only provide an opportunity to monitor one's peers and to accumulate social capital, but also add a social element to saving money that would not occur within a formal bank account.

Studies such as Anderson *et al.* (2009) or Besley *et al.* (1993) have analysed the question of ROSCA sustainability at the individual level and the conditions (or ROSCA design) under which default is more or less likely to occur. However, their predictions have not yet been tested empirically. All other existing field surveys of ROSCAs examine membership or group characteristics at one point in time. By definition, any groups that had collapsed

[1] ROSCAs have also been observed amongst ethnic minorities or immigrant groups in developed countries; see for example Bonnett (1981) or Summerfield (1995)

due to default or enforcement problems would have been omitted from such studies. As a result, we know relatively little about the lifespan of ROSCAs and their ability to deal with enforcement problems. However, this is of crucial importance: losing one's savings if another member defaults poses a potentially serious negative income shock to ROSCA members. Furthermore, there are examples of formal banks and microfinance institutions (MFIs) using the model of a ROSCA in their offerings (Handa and Kirton, 1999; El-Gamal *et al.,* 2014); a knowledge of the best structure or sociocultural makeup of ROSCAs might aid such institutions in designing their savings products in future.

This study presents a first empirical investigation into the sustainability of ROSCAs over time. I use a unique panel dataset of urban ROSCAs collected in Cotonou, Benin. After an initial field survey in 2004 where information was gathered on the characteristics of the membership and the groups themselves, a second survey round assessed the same groups in 2006. The analysis is therefore at the group level. The approach taken, whereby the same groups were surveyed twice, represents a significant improvement on the only other study (to my knowledge) that has empirically considered the sustainability of ROSCAs. Whilst Handa and Kirton's study (1999) acts as a first check on some of the relevant theories, the regressions contained therein suffer from two major problems. The dependent variable (whether or not a group has experienced enforcement problems) is based on retrospective information, so (i) any groups that have collapsed were, by implication, omitted from the study – it is shown below that often groups that have experienced problems collapsed - therefore the sample is biased toward those groups better equipped to deal with enforcement problems. And (ii), the use of retrospective information as the dependent variable leads to concerns over endogeneity; it might well have been the case that groups which experienced enforcement problems made changes to their institutional design or membership in an attempt to avoid such problems again. The present study avoids both issues.

The data paints a picture of inherent instability amongst Beninese ROSCAs – in the two years between the surveys, around one in three groups had experienced enforcement problems; the majority of these collapsed entirely because of this. Anderson *et al.* (2009) describe two potential causes of enforcement problems – a member who has received the pot and ceases to make further payments, and a member who receives an unfavourable rank in the allocation and fails to make any payments to the group at all. Using data from the 2004 wave of the survey, this study seeks to assess how various group characteristics can influence the incentives for members to default or miss payments and thus on the likelihood

that the group as a whole experienced such an occurrence in the following two years. The empirical analysis tests the main theories outlined in the theoretical literature - with regards to the effect of the order of pot allocation, group size, frequency of payments, level of social capital amongst members and the screening of new members - on the likelihood that a group experienced enforcement problems. The results of a Probit analysis finds support for the theory that ROSCAs can be designed in order to minimise the likelihood of experiencing enforcement problems occurring; groups that fix the allocation of funds based on members' need, as opposed to a random draw, were less likely to experience repayment problems, as were groups with written rules and those that met less frequently. The results also suggest that those groups led by a sole president, as opposed to a committee of members, were also significantly less likely to see problems occur. A number of proxies are considered in order to capture the level of social capital existing within groups and a number of interesting results are found, that add weight to the notion that social ties are key to enforcing ROSCAs. In particular, those groups started amongst family members appeared less likely to have experienced enforcement problems than those started amongst colleagues or friends. These results echo the findings of, for example, Karlan's study (2007) on social connections within group banks in Peru. A further way in which the risk of enforcement problems can be minimised is through the screening of new members. Thus, parallels are drawn to the work of, for example, Ghatak (1999), who showed that higher repayment rates might be possible where group lenders were able to choose new borrowers. The screening role might fall upon the president or ruling committee (or the group may collectively decide) and a variety of different screening devices (such as surveying new members) can be used. However, the results show no robust evidence that the screening functions of a group significantly impact the likelihood that they experience enforcement problems. Evidence is also presented that as groups get older or more experienced, the effect of the aforementioned features on the likelihood of experiencing enforcement problems diminishes, suggesting that repeated interactions in the ROSCA setting can help to enforce the bonds of trust and reciprocity amongst members.

The remainder of this chapter is organised as follows. Section 2.2 discusses the basic functioning of a ROSCA and introduces the Beninese sample, providing descriptive statistics of the key variables that will be tested in the empirical section. In section 2.3, the theory and intuitions shaping the incentives for a member to remain true to the ROSCA are discussed in detail. A number of conjectures are formed that are tested empirically in section 2.4. Section 2.5 considers some extensions, specifically asking if ROSCAs become

less likely to experience enforcement problems following repeated successful cycles of interaction. Section 2.6 concludes.

2.2 The Beninese Context and Descriptive statistics

2.2.1 Savings in Benin

ROSCAs are prevalent across the developing world, offering a means to save for those with limited access to formal alternatives, either due to missing or incomplete markets or because of high transaction costs. Helms *et al.* (2005) note that whilst there is a "deeply entrenched tradition" with regards saving in Benin, the majority of deposits are held informally, for example in ROSCAs or with itinerant bankers. Gracia (2000) found that around 81-89% of Beninese were using ROSCAs in a study carried out in the year 2000 and Helms *et al.* (ibid.) suggested that just 10% of deposits at this time were held in the formal banking system.

There are a number of reasons why such savings habits and trends exist in Benin, or at least, did in the mid-2000s. Firstly, the three state-run banks crashed in the late 1980s and many saw their savings frozen for a time. This led to mistrust in the banking system for those who remembered the crash (Helms *et al.*, *ibid.*). Secondly, banking coverage was poor and uneven: In 2004, Benin had 9 formal banks with a combined total of just 42 braches.[2] Including NGOs, Savings and Credit Co-operatives and the *Caisse Nationale d'Epargne* (CNE; the savings arm of the Post Office), there were still just 702 legally recognized points of service in the country at the time. Yet the majority of the country's formal savings were held in banks: some 83% as of 2004, with the remainder in the CNE or MFIs. It is also the case that many individuals enjoy the convenience and flexibility of informal savings. Thirdly, banks often impose restrictions on the opening or use of their services: for example, of the nine banks present in Benin in 2004, all required depositors to maintain a minimum balance, five required formal ID, three imposed an opening fee and a further three provided zero interest (Helms *et al.*, *ibid*). Moreover, the convenience of a ROSCA that might meet in one's locality, cannot usually be matched by having to travel to a bank branch which might not be located close by. Related to the mistrust of bankers, when individuals know the other members of a ROSCA, then they have peace of mind over

[2] However, even when formal alternatives are available, there is evidence that many individuals still choose to keep savings in a ROSCA (Gugerty, 2007; Dupas *et al.*, 2012).

who holds their deposits until they receive the pot, unlike when saving in a bank.

2.2.2 ROSCAs in Benin

ROSCAs consist of a group of people, *N*, who meet regularly to contribute a sum of money to a central 'pot' which is distributed to each member in turn. A 'cycle' is completed once every member has received the pot once. Beyond this basic framework, the groups vary across many dimensions such as size, the frequency of meetings, the order in which the pot is allocated and ruling structure. Their popularity is despite considerable drawbacks – the arrangements are less flexible compared to saving alone and no interest is earned on savings. If receipt of the ROSCA pot is viewed as a loan (which is a plausible comparison for the first half of members to receive it), then the absence of interest rates is further pronounced, i.e. in formal credit markets, interest rates are often used as a screening device to overcome the problem of adverse selection (Stiglitz and Weiss, 1981); this is obviously not possible within a ROSCA.[3]

Often it has been assumed or observed that ROSCAs are used to finance the purchase of a durable good (e.g. Besley *et al*, 1993), though Basu (2011) and Gugerty (2007) show how they might also be used as commitment devices for individuals with time-inconsistent preferences. As detailed by Dagnelie and LeMay-Boucher (2012), who use the same survey as this study, 49% of individuals surveyed in Benin used the pot for investments in their business, compared to just 14% who wished to acquire a durable good.[4] In terms of the motive for joining (i.e. regardless of the intended use of the pot, why an individual chose a ROSCA over saving in autarky), the most popular stated in the sample was due to commitment and self-control problems when saving by one's self.

The data for this study was collected during two waves of household surveys in 2004 and 2006 in Cotonou, Benin.[5] Surveys were carried out in two of the poorest areas of the city, namely Vossa and Enagnon. No banks, non-governmental organisations (NGOs) or other formal institutions offered saving facilities in these regions at the time. 497 households were selected in the first wave of the survey, according to a random process: 110 in Vossa, 387 in Enagnon. Within the 497 households, information on 182 active ROSCAs was

[3] Of course this disadvantage to the group becomes a significant advantage to those individuals who in effect receive an interest-free loan

[4] Other reasons given include for renovations or building a house (18%), school fees (7%) or to cover a debt (5%).

[5] The survey methodology is described in Appendix A

uncovered (when an individual responded that he or she was a member of a ROSCA, a questionnaire on the characteristics of the group was administered). 97 of these ROSCAs were surveyed again in 2006. Whilst the enumerators made every effort to track down members of all 182 groups in the second wave of the survey, it was not always possible to do so: many of the previously surveyed individuals may have left their family home or left the area entirely.[6] Therefore there is an attrition rate of 46.7%. However, aside from the reduced sample size, this only becomes a major issue if attrition is systematic along some of the variables of interest. Appendix B shows a test of equivalence of means between those groups surveyed in both 2004 and 2006 against those only sampled in the latter wave. The results of this show that only four out of the 25 variables (defined below) used in the empirical analysis had significantly different values between the groups surveyed in just 2004 and those surveyed in both waves. Indeed, the majority of variables were, on average, of very similar magnitude. As a further check, a Probit regression was run (not shown), where the dependent variable is a dummy equal to 1 if the group was surveyed in both waves. It appeared that very few variables of interest were significantly correlated with the likelihood of being surveyed again. Thus the likelihood of (not) being surveyed again appears to be down to mere coincidence, or at least circumstances not resulting from the features of the ROSCA.

The empirical analysis, which is presented below in section 2.4, considers the likelihood of a ROSCA experiencing enforcement problems during the 2004-06 period, based on its initial characteristics, measured in 2004. Table 2.1 thus presents summary statistics of the key characteristics of the ROSCAs in our sample measured in the original 2004 wave of the survey. It also includes a test for equivalence of means between groups that did not experience enforcement problems between 2004 and 2006 (64) and those that did (33). 'Enforcement problems' denotes a case where a group had experienced at least one case of non-payment between 2004 and 2006. The average ROSCA comprised around 33 members, although this ranged from 4 to 175. Just under half of the groups in this survey had remained the same size since inception (42 of 97). In 2004, each group had completed an average of over 8 cycles, although at the time of the survey, some had not yet completed their first. It appears that the subset of groups that experienced enforcement problems were on average significantly smaller, had completed fewer cycles and existed for a shorter time.

[6] However, there were no cases where the correct individual was located and they refused to be re-surveyed.

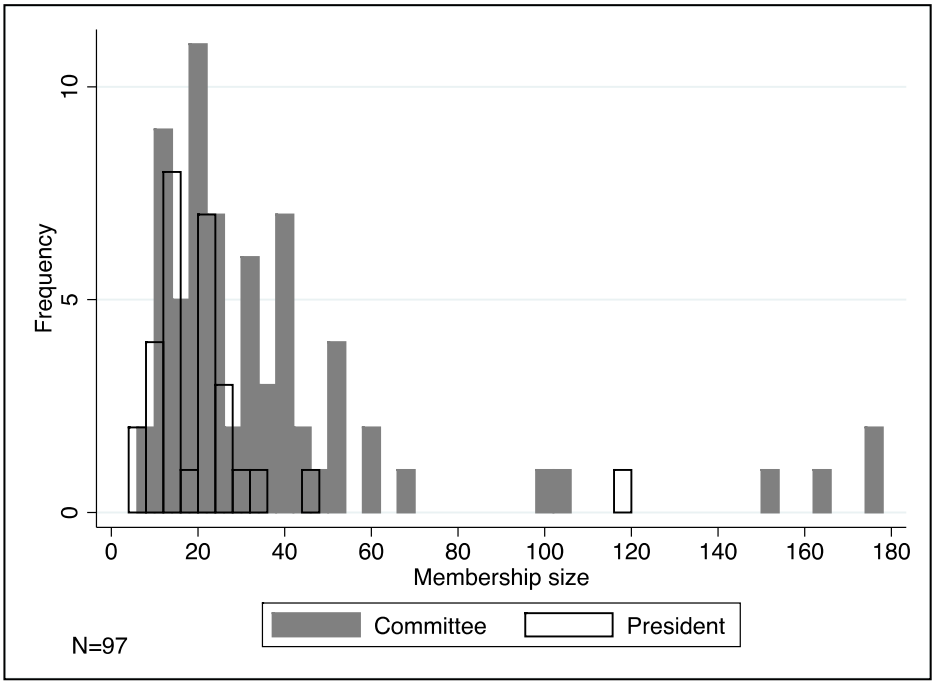
A key variable that could affect the incentives to default is the order in which the pot is allocated; this has duly attracted attention in the literature (see for example, Besley *et al.*, 1993; Anderson *et al*., 2009; Basu, 2011). Broadly speaking, this is done according to a fixed order, random draw (either at the start of a cycle or repeated at each meeting), or the pot is awarded to the highest bidder. Some 32% of the sample had a fixed order, with the remainder allocating the pot according to a random draw; no 'bidding' ROSCAs were observed in Cotonou. Of those groups that allocated the pot according to a fixed process, this was determined by i) the seniority of members (5 groups); ii) members in need of the money (13); iii) the regularity of attendance (6); iv) the appropriateness of a request (4) and v) the number of members (2). The variable *Need* is a dummy taking the value one if the pot was allocated according to criteria (ii). This might represent a situation where a member requests a specific rank in the cycle in order to cover some unforeseen expense, or it might be the case that members prefer to receive the funds at a certain date to cover regular lumpy expenses such as purchasing stock for a business or paying school fees etc. The variable *Other fixed* accounts for those groups setting the order of pot reception according to one of the other criteria listed.

Table 2.1 *Summary Statistics and test of equivalence of means between ROSCAs experiencing enforcement problems between 2004 and 2006 and those that did not*

	Full Sample				No Problems		Problems		Diff.
Variable	*Mean*	*Std. Dev.*	*Min*	*Max*	*Mean*	*S.E.*	*Mean*	*S.E.*	*Mean*
Experienced enforcement problems	0.340	0.476	0	1					
Membership size	33.175	33.930	4	175	37.46875	5.0239	24.84848	2.238	12.620*
No. of cycles completed	8.639	21.165	0.15	131	11.183	3.205	3.188	0.435	7.995*
Duration of existence (months)	67.309	93.594	1	480	79.063	13.399	44.515	9.289	34.547*
Only President	0.299	0.460	0	1	0.313	0.058	0.273	0.079	0.040
President/ committee paid	0.227	0.421	0	1	0.266	0.056	0.152	0.063	0.114
Random	0.680	0.469	0	1	0.703	0.058	0.636	0.085	0.067
Need	0.134	0.342	0	1	0.172	0.048	0.061	0.042	0.111
Other Fixed	0.186	0.391	0	1	0.125	0.042	0.303	0.081	0.082**
Written Rules	0.598	0.493	0	1	0.672	0.059	0.455	0.088	0.217 **
Monthly meetings	0.371	0.486	0	1	0.391	0.062	0.333	0.083	-0.057
Weekly meetings	0.351	0.480	0	1	0.313	0.058	0.424	0.087	0.103
More severe sanctions on delinquent member	0.536	0.501	0	1	0.875	0.042	0.727	0.079	0.148*
Pot size (1000's of CFA)	123.288	166.631	0.8	1000	132.470	24.027	105.479	17.461	26.992
Single ethnicity	0.237	0.428	0	1	0.219	0.052	0.273	0.079	-0.054
Only men	0.237	0.428	0	1	0.234	0.053	0.242	0.076	-0.008
Only woman	0.155	0.363	0	1	0.156	0.046	0.152	0.063	0.005
Survived Past Problems	0.206	0.407	0	1	0.063	0.030	0.485	0.088	-0.422***
Started amongst:									
Friends	0.433	0.498	0	1	0.344	0.060	0.606	0.086	-0.262**
Family	0.082	0.277	0	1	0.094	0.037	0.061	0.042	0.033
Members of same trade	0.144	0.353	0	1	0.156	0.046	0.121	0.058	0.035
Neighbours	0.186	0.391	0	1	0.203	0.051	0.152	0.063	0.052
Members of another group	0.072	0.260	0	1	0.094	0.037	0.030	0.030	0.064
Other	0.052	0.222	0	1	0.063	0.031	0.030	0.030	0.032
President/Committee decides	0.650	0.480	0	1	0.641	0.061	0.667	0.083	-0.026
New members must be known	0.608	0.491	0	1	0.609	0.062	0.606	0.086	0.003
Survey on new members	0.794	0.407	0	1	0.828	0.048	0.727	0.079	0.101
Known + Survey	0.464	0.501	0	1	0.500	0.063	0.394	0.087	0.106
Other conditions?	0.784	0.414	0	1	0.766	0.053	0.818	0.068	-0.053
	N = 97				*N = 64*		*N = 33*		

A lone president, as opposed to a committee of members, ran 30% of groups. The average size of a group run by a president (22) was significantly smaller than that overseen by a committee (38). Figure 2.1 highlights that larger groups were, aside from one outlier, more often ruled by a committee of members than a sole president. This might well be taken as a reflection of the greater workload for a sole president in larger groups. A committee of members can help spread the workload of the day-to-day running of the group such as organising meetings, screening new members etc.

Figure 2.1 *Choice of ROSCA ruling structure by size*



A significantly larger proportion of groups that did not experience enforcement problems had written rules (67%), compared to those that did (46%); around 60% of the full sample had rules. ROSCAs that had experienced problems pre-2004 appear much more likely to see a recurrence between 2004 and 2006 (denoted by the variable *experienced past problems)*. Almost half (49%; 16 of 33) of the groups experiencing problems between 2004 and 2006 had also done so in the past. Of those 16 experiencing a recurrence, 7 were able to overcome them and survive. However, of the 17 groups experiencing problems between 2004 and 2006 for the first time, only 3 were able to survive. This suggests that for some groups, enforcement problems are a somewhat regular occurrence but that they have devised means to cope with it.

Despite the substantial threat of social sanctions on members who miss payments, many groups also imposed more traditional forms of punishment on delinquents. During the survey, interviewees were asked about the punishments imposed in their ROSCAs, namely 'What are the sanctions imposed against a member who fails to pay a contribution?' A number of sanctions were used and the variable *More severe sanctions on delinquent member* represents the fraction of groups imposing at least one of those sanctions deemed to be more severe. These included '[the delinquent] receives a less favorable rank in the next cycle', '[the delinquent] is penalized at the breakdown of cash', '[the delinquent] receives his own contributions back but does not receive the pot at his turn', 'a fine must be paid,' '[the delinquent] does not receive the pot at his turn', 'property seized' or 'the member is excluded'. Conversely, those classified as less severe are 'no penalty', 'the person receiving the pot in this period need not contribute to the pot when it is the delinquent's turn to receive the pot' and [the delinquent] is given a period in which to pay.' The majority of groups imposed 'more severe sanctions' upon a delinquent.

The pot size ranges from 800CFA (USD1.50) to 1,000,000CFA (USD1862), with an average of 123,000CFA (USD230).[7] 35% of groups met once per week and 37% of the groups met just once a month. One met daily whilst others met every two weeks (8%), every ten days (3%), or 2-3 times per week (16%).

Whilst inherently difficult to measure, this study uses a number of proxies in an attempt to capture the level of social capital that exists between ROSCA members. Notably, groups were asked about the pre-existing social connections, or similarities, that existed at inception - the majority were started amongst friends (43%), family (8%) or neighbours (19%); around 24% were exclusively for members of the same ethnic group (denoted by the variable *Single Ethnicity*) and around 40% of groups were restricted to either only males or only females (denoted by the variables *Only men* and *Only Women*). 61 % of those groups that experienced enforcement problems between 2004 and 2006 were started amongst friends, compared to only 34% of those not experiencing problems. This difference is significant at the 5% level. Table 2.1 also summarises the information with regards the screening roles performed by ROSCAs in the sample. A majority of groups surveyed new members or stipulated that entrants must be known to the current membership (61% and 79% respectively); just under half of the sample imposed both restrictions and only 6% imposed neither restriction. Either the president or ruling

[7] June 2004 exchange rate of 537CFA to US$1 used

committee were entrusted with the decision over whether to admit new members in 65% of the groups in our sample (denoted by the variable *President/Committee decides*), with the remaining groups deciding as a whole. Individuals were also asked whether there were any additional constraints imposed upon people wishing to join. For example, it is often stipulated that new members must have a certain occupation. The variable *Other Conditions* signifies that around 78% of groups did impose such constraints.[8] 60 % of the interviewees in the Beninese data stated that, given the choice, they preferred to receive the pot towards the end of the cycle, whilst fewer than one in four preferred to receive the pot at the beginning

2.3 Institutional Design and Social Capital

This section provides a discussion of how the incentives to remain true to the ROSCA arrangement can be affected by various aspects of the group design, drawing from relevant literature where appropriate. I specifically focus on (i) the institutional design of ROSCAs (including the screening effort performed on new entrants) and (ii) the level of social capital that exists between members. The discussion is used to form conjectures which will be tested in section 2.4. Given that the literature on ROSCA design and empirical evidence on repayment rates is quite limited, some comparisons are drawn to joint-liability microfinance groups, which are comparable in nature but have been studied in relatively more detail.

In the absence of a well-designed incentive mechanism, there will be a temptation for the utility maximising ROSCA member to take the collective savings and default on further payments.[9] Besley *et al.* (1993) formalise this argument to show that in order for a member to stick to the agreement, there must be some cost of default, *K*, that outweighs the benefits.

Specifically,

$$K \geq \left(\frac{(N-r)}{N}\right) B \left[\frac{v(1,y)-v(1,c)}{y-c}\right] \qquad [1]$$

[8] Common answers were that new members must 'have an income generating job' or simply 'have good morals'.

[9] Or, for someone who has been allocated a later rank, the temptation would be to cease making payments before they receive the pot.

where N is the group size, r is an individuals' rank, B is the size of the pot (assumed, in this case to be used for consumption of a durable good), y an individuals' income and c nondurable consumption.[10] The second argument on the right hand side of [1] is thus an individuals' utility given that the durable good is owned.[11] This basic framework highlights that the incentives to default can be affected by both the costs (K) of doing so, and the benefits. The latter can be influenced by size of the pot (B), group size (N) or an individuals' rank (r), which itself is influenced by the means by which ranks are allocated. Whilst not modelled explicitly in Besley *et al.*, the costs of default, K, might depend on the availability of formal savings alternatives (zero for individuals in the survey regions of Cotonou), the availability of other informal savings devices, the potential loss of reputation, or other sanctions imposed by groups on delinquent members (e.g. property might be seized, or the member might be excluded from joining future informal groups in the area). It is likely that the cost of default will be higher where groups are formed amongst those with stronger pre-existing social connections such as family.

Thus there is significant potential for ROSCAs to be designed in such a way that minimises the individual's incentive to default. A plausible assumption when studying ROSCAs is that the same set of incentives will apply to all members and thus the group as a whole. One can justify this by arguing that members are relatively homogenous individuals (for example having the same trade or coming from the same neighbourhood) who are able to commit to the same saving patterns over time. Therefore, an analysis of the characteristics of the group itself should prove insightful.

2.3.1 Institutional Design

The allocation of ranks plays an important role in influencing the incentives to default or remain true to the group. As a simple illustration of how these incentives differ according to how the pot is allocated, consider a ROSCA with N=10 in its first cycle. Under a fixed allocation, the tenth recipient will feel tempted to cease making payments as they are only doing just as well as saving alone for ten periods, whilst at the same time risking their savings in the hands of other members and possibly following a sub-optimal savings pattern

[10] Adapted from page 806 of Besley *et al.* (1993).
[11] It is not important here that purchase of a durable good was not the sole reason given for joining ROSCAs in our sample. We can consider the utility for whatever a person chooses to spend the money on as analogous to that gained from being able to purchase a durable good.

(Anderson *et al*., 2009).[12] If the fixed allocation is kept constant across cycles, they must also wait until the final meeting of the second cycle to receive the pot a second time. Yet in a random-order ROSCA, there is a 90% chance that this last-placed individual will receive the pot in an earlier round of the next cycle.

A member's rank, *r*, is one of the key parameters influencing sustainability in the model of Anderson *et al.* (2009). They illustrate that whilst a majority of members would prefer the ranks to be allocated according to a random draw, this will actually tend to increase the likelihood of experiencing enforcement problems. This result, however, relies on a common assumption in this literature that assumes ROSCAs are primarily used to save for the purchase of indivisible durable goods. Thus, receiving the pot earlier in the cycle allows individuals to make the desired indivisible expenses sooner than if they had saved alone. However, such an assumption does not always hold. Gugerty (2007), using Kenyan data, and Dagnelie and LeMay-Boucher (2012), using the same Beninese sample as in this study, provide evidence that ROSCAs can be used as commitment devices. Instead of desiring an early place in the cycle, members would rather opt for later ranks in order to discipline themselves to save. It is worth noting that 60 % of the ROSCA members in our survey stated that given the choice, they would prefer to be allocated a later rank and fewer than one in four preferred to receive the pot at the beginning. Given this, it is difficult to say whether fixed or random ROSCAs are more likely to be sustainable. However, it is notable that some fixed order groups allocate ranks by accommodating, to some extent, their members' desires and needs. Such groups might be able to allocate ranks so as to minimise problems relating to payment. Specifically, a member facing financial difficulties, such as an unexpected income shock, might be allocated the pot at an earlier meeting. Similarly, members wishing for a late rank due to some commitment motive could be accommodated. Handa and Kirton (1999) note that the presidents in their Jamaican sample also offer similar arrangements.

Conjecture 1*: Given that a majority of individuals in the sample use ROSCAs as a commitment device, it is not expected that the order of pot allocation (random/fixed) will play a significant role in the likelihood of experiencing enforcement problems. However, fixed order ROSCAs which accommodate their members' needs may minimise the likelihood of problems occurring.*

[12]That said, often individuals might actually prefer to place savings in a ROSCA than to keep the money at home: the risks here are also substantial (fire, theft, claims from family members etc.)

The supervision of the group is another feature of ROSCA design that has received attention. In our Beninese sample, ROSCAs are overseen by either a president or a committee of members (ruling committee). The costs of missing payments or defaulting (*K*) can be significantly affected by how effectively the group is monitored or run and by how credibly those in charge reveal and enforce the threats of punishment to any potential delinquents. Paxton *et al.* (2000), for example, showed that a strong leader led to higher repayment rates of group loans in Burkina Faso.

Anecdotal field evidence from a pilot survey indicated that given a set of similar group characteristics, president-led ROSCAs are generally more tightly run than those with a ruling committee. Presidents also appeared more involved in informally monitoring the activities of their members. This can be partly explained by the fact that a larger share of president-led ROSCAs in our sample are remunerated by the group for organising meetings and providing supervision. Around half the presidents were paid, but only one in ten committees received remuneration for their role. The incentives for the rulers, who are also a part of the ROSCA, to carry out their duties with greater diligence might be enhanced by the offer of remuneration. Handa and Kirton (1999) found that payments to the leader significantly increased the sustainability of ROSCAs, although close to 90% of presidents were paid in their Jamaican sample. Furthermore, a number of presidents (8 out of the 29) are founding members, whereas this is less frequently the case for ruling committees. Founding members are often well respected figures who can act with greater authority. They are also more likely to have a good understanding of the actions required to avoid and potentially solve problems.

In some fixed order ROSCAs, the decision over the order of pot reception is often used as a selection tool, analogous to a credit rating i.e. the first to receive the pot is allocated the most credit.[13] A new member would likely be allocated a less favourable rank and upon successful completion of one or more cycles, his creditworthiness might improve and he might be allocated a more advantageous rank in future cycles.[14] Groups where either the president or the ruling committees can decide on this matter could be less likely to

[13] See, for example, Van den Brink and Chavas (1997) and Handa and Kirton (1999).
[14] Biggart (2001) notes that this signal might emerge not only as a result of the new member having had an entire cycle to prove himself reliable with regards repayment, but also by socialising with his peers.

experience enforcement problems. In line with the preceding argument, presidents do have the power to unilaterally decide the order more often than ruling committees.[15]

Conjecture 2*: Groups (i) overseen by a president rather than a committee, (ii) where the rulers receive remuneration, or (iii) where they can decide the order of pot allocation, are expected to have a lower likelihood of experiencing enforcement problems.*

The size of the pot of collective savings, *B*, is another variable that could influence the incentives for a member to make, or cease to make payments. For a given membership size, a larger pot provides a greater temptation for a potential delinquent member to default (either one who has just received the pot, or one who must make many [large] payments before receiving). Besley *et al.* (1993) postulate that a larger pot will likely create larger incentives to default (or rather, it is the ratio of default cost to pot size that matters [$\frac{K}{B}$ from equation 1]). One relevant counter argument is that a larger pot might actually provide greater incentive for those who are using the ROSCA as a commitment device to save even more, as the eventual sum received is greater.

Conjecture 3*: A larger pot increases the incentive for any one member to cease contributing once (s)he has received it and hence leads to a higher likelihood of the group experiencing enforcement problems*

The frequency of meetings can also impact upon the likelihood that a group experiences enforcement problems. Besley *et al*. (1993) postulate that, holding membership size constant, increasing the duration of the ROSCA can mitigate against the likelihood of default. Indeed, Handa and Kirton (1999) present some evidence to this effect; their study found that Jamaican ROSCAs meeting once per month were (albeit, only weakly significantly) less likely to experience problems than those meeting weekly. Payment frequency has also been considered in studies of MFIs. McIntosh (2008) saw increased repayment rates of FINCA clients in Uganda that met bi-weekly as opposed to those meeting weekly; Field and Pande (2008) argue that more frequent repayment provides microfinance clients with a chance to help form the habit of saving and that a rigid repayment schedule enabled clients to develop 'fiscal discipline', although they found no difference in repayment rates between clients on monthly and weekly repayment schedules.

[15] The president established the order in 57% of the 30 fixed order groups, the committee in 33% and 10% of groups decided as a whole.

On the contrary however, Armendariz and Murdoch (2005) and Feigenberg *et al.* (2013) noted higher default where repayment was less frequent in Bangladesh and India respectively. Thus it seems that the theory and majority of evidence points to a lower frequency of meetings leading to higher repayment rates and a lower likelihood of enforcement problems.

Conjecture 4*: For a given number of members, ROSCAs that meet less frequently are less likely to experience enforcement problems.*

In addition to structural features that vary across all groups (such as order of pot reception, frequency of meetings, role of the president etc.), there are some that are group-specific. For example, whether or not a group has written rules, or if they impose sanctions on delinquent members. When the ROSCA has written rules, this might add a sense of formality to the functioning of the group. Indeed, if members have agreed to adhere to a set of rules outlining, for example, when and where meetings take place, the sanctions against a defaulting member, or any of the other features outlined above, then they might have a greater sense of their obligations to the group, thus lowering the likelihood that the group experiences problems. Whilst evidence in this regard is scant, Wenner's (1995) study of 25 group credit programs in Costa Rica suggested that those with a 'written code' saw lower levels of delinquency, as did Zeller's study in Madagascar (1998).

Conjecture 5*: ROSCAs with written rules are less likely to experience enforcement problems.*

The various sanctions faced by delinquent members were outlined in section 2.2. To the extent that these present a meaningful punishment on potential delinquents, there would likely be a lower likelihood of a group experiencing enforcement problems, where more severe sanctions were in place.

Conjecture 6: *In addition to the threat of social sanctions, those groups imposing other sanctions upon delinquent members will be less likely to see enforcement problems.*

Whilst many ROSCAs in the sample did not vary in size, some did allow new members to join at the start of a new cycle. In this case, there might well be a degree of informational asymmetry with regards their reliability. It is therefore in the interests of the group to screen

new entrants and in the interests of individuals to signal their reliability. The intuition is analogous to the work of Ghatak (1999), who showed that where group lenders were able to select new borrowers, they would potentially be able to achieve higher repayment rates. Often restrictions are placed upon new members; some groups are composed entirely of one ethnicity, one gender or of individuals from one company or trade. It stands to reason that where new members are more effectively screened, the group will be able to select those that are, *ex ante*, considered to be more reliable.

Conjecture 7: *ROSCAs that effectively screen new entrants are less likely to experience enforcement problems.*

2.3.2 Social Capital

Formal credit institutions require that lenders front some form of collateral. However, this is not usually the case with ROSCAs. In practice, members might incur fines or have their possessions seized if they miss payments, but in the ROSCA context, the idea of social sanctions arising from the social capital existing amongst members, are perhaps more important.[16] Putnam *et al.* (1993) argue that this can be viewed as a kind of collateral, especially in cases where no physical assets are available. Karlan (2007), for example, found evidence that strong social connections amongst group members lead to higher repayment rates in a group banking scheme in Peru, although Wydick (1999) found little evidence of this in Guatemala.

As such, the fronting of so-called *social collateral* (Besley *et al.* 1993) helps to overcome the problem of adverse selection by acting as both a screening device on new members and an enforcement mechanism amongst existing participants. The loss of reputation arising from non-payment can be viewed as a 'social sanction'; such punishment becomes less desirable when higher levels of pre-existing social capital exist between members and it is conceivable that individuals would go to great lengths to avoid a bad reputation amongst family, neighbours, friends, colleagues or business partners. The importance of social sanctions will be greater still in those areas where access to formal credit markets is more restricted; if informal institutions that screen entrants purely on reputation are the prevalent savings option in a region as is the case for our Beninese sample, then ROSCAs provide an attractive solution for individuals wishing to smooth their consumption or commit to save.

[16] Putnam *et al.* (1993:167) define social capital as '...features of social organisation such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions.'

The threat of losing one's social capital is therefore what could make up the bulk of the cost of default, *K*, in the model above.

Whilst social capital may be pre-existing amongst members and as such facilitate the formation of a ROSCA (they are often started amongst members of the same neighbourhood, workplace, family, or religious group), it is also possible that this could be accrued over time; trust and reciprocity can emerge as a result of repeated interactions, even between strangers. In ROSCAs, interactions between members are not only a result of financial motivations – groups often perform an important social function, for example meeting together for meals, or members doing business with one another. Such interactions can only serve to enhance cohesion amongst the group.

Conjecture 8: *Where a higher level of social capital exists amongst members, there is a lower likelihood that the ROSCA will experience enforcement problems.*

2.3.3 Monitoring and Peer Pressure

Repeated successful interactions in a ROSCA might go some way to circumventing the likelihood of enforcement problems in that it provides a signal of trustworthiness, yet there is still a need for members to monitor each other. This is especially true for those members who have already contributed to, but not yet received, the pot; they are in effect seeking to protect their investment. Effective monitoring might lead to earlier and easier detection of problems, allowing the group to better deal with a member who is facing difficulties in making his or her payment. Monitoring is possible due to the social connectedness of group membership: where members interact outwith the ROSCA setting (perhaps as part of the same business, neighbourhood or church), they are able to keep an eye on their colleagues. However, the ability of members to monitor each other is inherently dependent on numerous variables such as group design, size, or the frequency of meetings. Chiteji (2002) considers that the ability of members to enforce the arrangement depends upon characteristics such as innate ability to police each other (peer monitoring), the value placed upon reputation, the awareness of another's reputation (*ex-ante* screening, or effort exerted to monitor the actions of another) and the existence of pre-existing relationships between group members that can be used to extract information with regards their reliability or integrity. Chiteji (2002) envisages that very large groups would involve monitoring costs outweighing the expected benefits from ROSCA participation and in such cases it would

not make sense to form one. Besley *et al.* (1993) also postulated that default risk could be overcome by *ex ante* reducing group size.

Kandel and Lazear's (1992) examination of the roles of peer pressure and monitoring within partnerships (firms where profits are shared amongst workers) and the insights presented are also quite relevant: the fundamental incentives in such firms are analogous to those in a ROSCA: a workers' choice of effort in a partnership affects the level of profit received by all other staff, whilst ROSCA participants' choice of whether or not to make their contribution dictates how much (if anything) all the members will receive. They note that peer pressure might induce a worker to feel 'shame' (the disutility received when one's unsatisfactory efforts are noticed by colleagues), or 'guilt' (the disutility experienced by a worker whose actions negatively affect the wellbeing of his colleagues). It is likely that in ROSCAs, where social connectedness is often high, the latter feeling would only serve to reinforce the pressures already felt from the threat of social sanctions. The authors also note that peer pressure and monitoring becomes more difficult within larger groups.[17] Furthermore, Kandel and Lazear argue that feelings of guilt might induce loyalty, even when actions are unobservable. Given that the most cited reason for ROSCA membership in our dataset is that of self-control (Dagnelie and LeMay-Boucher, 2012), one can envisage the feeling of guilt as a mechanism that helps to further curb any behaviour that might reduce an individuals' ability to make payments (such as spending on frivolities)[18], therefore helping to reinforce group stability. Such feelings of guilt are likely to be stronger still where social connectedness is higher – i.e. amongst members of (for example) the same family, neighbourhood or ethnic group – a feature of many of the ROSCAs observed in Benin.

Recent work by Ahlin (2015) considered the role of size in a group lending context. His study uncovers a complementarity between group size and the level of social capital between members that suggests larger groups can achieve efficient outcomes, given a high enough level of local borrower information. Yet it is important to note that the author focuses on groups between 1 and 20 members; in our setting where the average ROSCA

[17] It can also be argued that peer pressure or monitoring exerted by each individual need not necessarily increase in larger groups (thus overcoming the problems cited in Kandel and Lazear (1992) and Chiteji (2002)). The mere fact that more people are aware of members' actions in a very large group may be sufficient deterrent against any threat of default. If ROSCAs meet together to make payments, an individuals' reputation may take a much greater hit when it is announced to a larger number of people that he is unreliable. As such, the severity of social sanctions may be high in a larger group.

[18] See Bonan *et al.* (2013) for a fuller discussion of the issues of preferences and frivolous spending amongst Beninese households.

size is 33 members (and can number well over 100), it is unlikely that local borrower information would be of sufficient quality to have the same effect as described in Ahlin (2015); i.e. the definition of a 'large' group can differ substantially between contexts.

Conjecture 9: *Enforcement problems are more likely to occur in larger groups due to higher monitoring costs and a lower level of social capital, on average, between members.*

2.4 Empirical Strategy and Results

This section tests each of the conjectures outlined above; Probit analysis is used to estimate the following model:

$$\Pr(\text{Enforcement Prob} = 1)_{i,04-06} = \alpha + \gamma \mathbf{X}_{i,04} + \theta \mathbf{S}_{i,04} + \varepsilon_{i,04} \quad [2]$$

where the dependent variable is a dummy variable equal to one if the ROSCA, *i*, experienced an enforcement problem between 2004 and 2006. **X** is a vector of characteristics related to the institutional design of the group, **S** is a vector of characteristics that proxy the level of social capital existing amongst members and ε_i is the group-specific error term. All of the right hand side variables are measured in 2004, allowing us to examine their effect on the likelihood of experiencing enforcement problems during the two years that followed. The results of Probit estimations are shown in Tables 2.2 and 2.3. Average Marginal Effects (AME) are shown.[19]

[19] Whilst it is often common to calculate and present marginal effects at fixed values of the independent variables such as the mean (MEM; i.e. at $x = \bar{x}$), this is inappropriate when there are dummy variables amongst the regressors, which is the case here. Bartus (2005) highlights that the means of dummy variables actually refers to non-existent observations. E.g. in the mean of *President* is 0.35, but to calculate the marginal effect of having *35% of a president* is not realistic. I therefore display the average marginal effects (AME) for each of the regressors i.e. the average of the marginal effect at each $x = x_i$.

Table 2.2 *Probit estimation*

Dependent Variable =1 if ROSCA experienced enforcement problems between 2004 and 2006

	1	2	3	4	5	6	7	8	9
Membership size	-0.009***	-0.009***	-0.009***	-0.009***	-0.007***	-0.008***	-0.008***	-0.008***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Number of Cycles	-0.012**	-0.016*	-0.012**	-0.012***	-0.015	-0.012**	-0.011***	-0.012***	-0.010***
Completed	(0.006)	(0.009)	(0.006)	(0.004)	(0.010)	(0.006)	(0.003)	(0.003)	(0.005)
Random	-0.004	0.112		-0.000	-0.021	-0.019	-0.017	-0.021	-0.016
	(0.076)	(0.075)		(0.075)	(0.078)	(0.074)	(0.072)	(0.072)	(0.074)
Need	-0.517***		-0.513***	-0.520***	-0.461***	-0.494***	-0.471***	-0.477***	-0.452***
	(0.108)		(0.115)	(0.109)	(0.101)	(0.103)	(0.101)	(0.101)	(0.101)
Other Fixed			0.004						
			(0.076)						
President	-0.268***	-0.225**	-0.268***	-0.288***		-0.268**		-0.271***	-0.260***
	(0.092)	(0.090)	(0.092)	(0.107)		(0.106)		(0.085)	(0.081)
Committee							0.265***		
							(0.100)		
Paid				0.030	-0.127	0.239**	0.912***		
				(0.091)	(0.080)	(0.099)	(0.130)		
*President*Paid*						-0.259**			
						(0.114)			
*Committee*Paid*							-0.922***		
							(0.133)		
Written Rules	-0.245***	-0.265***	-0.245***	-0.245***	-0.164**	-0.244***	-0.233***	-0.208***	-0.299***
	(0.073)	(0.072)	(0.073)	(0.073)	(0.070)	(0.071)	(0.070)	(0.066)	(0.077)
More severe sanctions on delinquent member								-0.134**	
								(0.066)	
Survived past	0.776***	0.630***	0.776***	0.773***	0.732***	0.753***	0.722***	0.754***	0.707***
Problems	(0.100)	(0.061)	(0.100)	(0.100)	(0.086)	(0.089)	(0.087)	(0.104)	(0.077)
Pot size (CFA 000's)	0.000	0.000*	0.000	0.000	0.000	0.000	0.000	0.000	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Pot size ^{2}									-2.66e-06**
									(1.19e-06)
Monthly Meetings	-0.148**	-0.188**	-0.148**	-0.145**	-0.166**	-0.162**	-0.156**	-0.155**	-0.223**
	(0.071)	(0.079)	(0.071)	(0.070)	(0.077)	(0.071)	(0.068)	(0.064)	(0.066)
Single Ethnicity	0.047	-0.008	0.047	0.049	0.002	0.034	0.022	0.021	0.045
	(0.072)	(0.080)	(0.072)	(0.072)	(0.072)	(0.071)	(0.070)	(0.072)	(0.070)
Only Men	-0.248***	-0.224**	-0.248***	-0.248***	-0.167*	-0.238***	-0.214**	-0.263***	-0.183**
	(0.095)	(0.091)	(0.095)	(0.096)	(0.089)	(0.091)	(0.091)	(0.089)	(0.089)
Only Women	-0.038	-0.066	-0.038	-0.032	-0.090	-0.018	-0.007	-0.026	0.027
	(0.084)	(0.096)	(0.084)	(0.085)	(0.095)	(0.082)	(0.080)	(0.084)	(0.075)
Pseudo – R^2	0.55	0.48	0.55	0.55	0.55	0.56	0.57	0.58	0.62
Observations	97	97	97	97	97	97	97	97	97

** p<0.1; ** p<0.05; *** p<0.01. Average marginal effects shown. Robust Standard errors in parentheses*

2.4.1 Institutional Design

Conjecture 1*: Given that a majority of individuals in the sample use ROSCAs as a commitment device, it is not expected that the order of pot allocation (random/fixed) plays a significant role in the likelihood of experiencing enforcement problems. However, fixed order ROSCAs which accommodate their members' needs may minimise the likelihood of problems occurring.*

Columns 1-3 of table 2.2 consider the effects of the order of pot allocation on the likelihood of experiencing problems. *Random* is a dummy variable taking the value 1 if this is decided according to a random process. *Need* is a dummy taking the value 1 if the allocation of the pot was done according to members' need for the money (discussed above). The reference category, in column 1, is a group where the order of pot allocation was fixed according to some other criteria such as seniority or regularity of attendance (listed in section 2.2 above). It is clear that those groups allocating the pot according to members' need were significantly less likely to have experienced enforcement problems than those allocating the pot to another criterion. The results in column 2 suggest that random groups do not appear to be any more or less likely to experience enforcement problems than *all* fixed groups, as the coefficient is not statistically significant. Column 3 changes the reference category to *random*, controlling for those groups fixing the order according to some other criteria. Again, the results suggest that groups allocating the pot according to *Need* are less likely (than random groups, this time) to have experienced enforcement problems. The average marginal effect of around -0.5 suggests that groups allocating the according to members' need were over 50% less likely to have experienced enforcement problems compared to those that fixed according to some other process or randomly allocated the pot. There appears to be no significant difference in the likelihood of enforcement problems between those allocating the pot randomly and any *Other fixed* process.

Conjecture 2*: Groups (i) overseen by a president rather than a committee, (ii) where the rulers receive remuneration, or (iii) where they can decide the order of pot allocation, are expected to have a lower likelihood of experiencing enforcement problems.*

The variable *President* is a dummy variable taking the value 1 if a single president oversees a group. Thus the reference group is one that is overseen by a committee of members. *President* is negative and statistically significant throughout the estimations presented. The

magnitude of the average marginal effects suggests that groups overseen by a president alone were between 23 and 27% less likely to experience enforcement problems compared to those ruled by a committee.

Also tested is the effect of groups having rulers that were paid for their work. The variable *paid* is a dummy equal to one if the ROSCA was led by a president or committee that was paid. When interacted with *President* (Col. 6), the negative and significant coefficient presents evidence that those groups where a president was paid were less likely to have experienced enforcement problems than those with an unpaid president. Col. 7 investigates the effect of a committee that is paid and a similar result emerges, though the effect is much stronger. That is, whilst it appears having a president in itself leads to a lower likelihood of enforcement problems, a group with a committee can lower the likelihood to a greater extent than those run by a president alone by providing remuneration to the leadership.

I also investigated whether presidents or committees having the power to determine the order of pot allocation affected the likelihood of enforcement problems occurring, however there was no significant effect.[20] Taken hand-in-hand with the above result that those groups allocating the funds according to members' need were least likely to have experienced enforcement problems, this suggests that it is often the decision criteria itself, and not the decision maker, that can minimise the probability of enforcement problems occurring in a ROSCA.

***Conjecture 3**: A larger pot increases the incentive for any one member to cease contributing once (s)he has received it and hence leads to a higher likelihood of the group experiencing enforcement problems*

In groups where the pot is larger, there should be a higher incentive for any one individual to cease contributing due to the larger monetary payoff, denoted *B* in the model described in equation [1] above. However, the estimations suggest limited evidence that the size of the pot is related to the probability of enforcement problems occurring. Whilst the variable *pot size*, is statistically significant at the 10% level in column 2 of table 2.2, this result is not robust to inclusion of other controls. When the square of *pot size* is included in column 9 of table 2.2, there is evidence of a quadratic relationship with the dependent variable. One explanation for this result might be that for small values, the pot is either not of high enough

[20] Results not shown.

value to default, or the group is very small and so monitoring is higher. For larger pot sizes, the monitoring effort might also be higher as members seek to protect their investment. Thus, I find only limited support for the predictions of the model in Besley *et al.* (1993) that a larger pot would increase the likelihood of default.

Conjecture 4*: For a given number of members, ROSCAs that meet less frequently are less likely to experience enforcement problems.*

Included in all specifications is a variable *Monthly Meetings*, which is a dummy equal to one if the group met just once a month, the reference category being those groups meeting more regularly.[21] Across all specifications in table 2.2, it appears that those groups meeting monthly were indeed less likely to have experienced enforcement problems than those meeting more frequently. Those groups meeting monthly had a significantly higher average pot size (209 CFA 000's) than those meeting more frequently (73 CFA 000's) and the correlation between the two variables is quite high (0.40). *Monthly Meetings* remains significant following the exclusion of the pot size control.[22]

Conjecture 5*: ROSCAs with written rules are less likely to experience enforcement problems.*

The variable *Rules* is a dummy taking the value of 1 if groups had written rules. This variable is one of the most strongly related to the likelihood of having experienced enforcement problems. Across specifications, the coefficient is consistently around -0.2 and significant at the 1% level, suggesting that the likelihood of enforcement problems occurring in those groups with rules was around 20% lower, on average, than those without.

Conjecture 6: *In addition to the threat of social sanctions, those groups imposing other sanctions upon delinquent members will be less likely to see enforcement problems.*

Column 8 of table 2.2 introduce a control for the sanctions imposed upon delinquent members. These sanctions are known by group members and in the vast majority of cases included in the written rules of a group. Specifically, this variable is a dummy taking the

[21] The breakdown for the frequency of meeting is as follows: monthly (36 groups); twice per month (8); every ten days (3); weekly (34); every 5 days (12); twice per week (2); every 2 days (1) and daily (1).
[22] This result is not shown for the sake of brevity.

value 1 if the group imposed severe sanctions on a delinquent member.[23] The results suggest that those groups threatening more severe sanctions on a potential delinquent were around 13% less likely to have experienced enforcement problems on average.

Conjecture 7: *ROSCAs that screen new entrants are less likely to experience enforcement problems.*

In columns 1-7 of table 2.3, attention turns to the screening roles performed by ROSCAs observed in the sample. Taking groups where the entire membership decides as a reference, a dummy variable is included, taking the value of one when the president or committee makes the decision over whether to allow new entrants (Denoted by the variable *President/Committee decides*), however this is not statistically significant. There is also no statistically significant relationship when this is interacted with the 'paid' variable (column 2). Columns 3-7 include the variables *Known*, *Survey* and *Other Conditions* (as defined in section 2.2 above). Again, however there appears to be no statistically significant relationship with the likelihood of enforcement problems occurring, although for the most part these variables do appear to have the expected sign. Thus, at least in the Beninese context, there is little evidence that the screening measures performed on new entrants act as an effective deterrent to potential delinquents.[24]

2.4.2 Social Capital

Conjecture 8: *Where a higher level of social capital exists amongst members, there is a lower likelihood that the ROSCA will experience enforcement problems.*

I next examine the pre-existing social connections between ROSCA members. Table 2.1 highlighted that groups in the sample are most frequently started amongst friends, but can also be formed with neighbours, family, members of the same trade, or members of another group. Dummy variables are included in column 8 for *n-1* of these categories, taking *Started amongst family* as the reference group, as such groups are likely to have the highest level of pre-existing social capital.[25]

[23] Those definition of sanctions as 'severe' is outlined in section 2.2 above.
[24] A Wald test, not shown, suggested that the screening variables are also jointly insignificant.
[25] A counter argument here is that groups containing many members of the same family might see the sanctioning / enforcement mechanisms undermined due to a reluctance to punish each other. (Sharma and Zeller, 1997).

Table 2.3 *Probit estimation*

Dependent Variable =1 if ROSCA experienced enforcement problems between 2004 and 2006								
	1	2	3	4	5	6	7	8
Membership size	-0.008***	-0.007***	-0.008***	-0.008***	-0.009***	-0.008***	-0.008***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of Cycles	-0.014	-0.012	-0.014	-0.014	-0.015	-0.012***	-0.012**	-0.013*
Completed	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.005)	(0.005)	(0.007)
President	-0.279***	-0.267***	-0.288***	-0.277***	-0.306***	-0.282***	-0.290***	-0.263***
	(0.090)	(0.086)	(0.093)	(0.090)	(0.094)	(0.088)	(0.091)	(0.098)
Random	0.004	-0.017	0.014	0.003	0.022	0.005	0.014	-0.011
	(0.077)	(0.074)	(0.077)	(0.076)	(0.075)	(0.077)	(0.075)	(0.078)
Need	-0.494***	-0.430***	-0.487***	-0.494***	-0.501***	-0.479***	-0.475***	-0.474***
	(0.109)	(0.092)	(0.108)	(0.109)	(0.107)	(0.103)	(0.103)	(0.126)
Written Rules	-0.243***	-0.232***	-0.241***	-0.243***	-0.236***	-0.250***	-0.248***	-0.201**
	(0.073)	(0.068)	(0.072)	(0.073)	(0.074)	(0.078)	(0.078)	(0.083)
Survived Past	0.768***	0.712***	0.785***	0.770***	0.781***	0.752***	0.769***	0.783***
Problems	(0.112)	(0.093)	(0.122)	(0.116)	(0.118)	(0.099)	(0.116)	(0.131)
Pot size (CFA 000's)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Monthly Meetings	-0.157**	-0.183***	-0.170**	-0.159**	-0.149**	-0.158**	-0.170**	-0.197***
	(0.072)	(0.069)	(0.075)	(0.073)	(0.072)	(0.072)	(0.076)	(0.072)
Single Ethnicity	0.056	0.020	0.043	0.055	0.057	0.051	0.040	0.111
	(0.073)	(0.069)	(0.077)	(0.074)	(0.071)	(0.074)	(0.078)	(0.083)
Only Men	-0.230**	-0.170**	-0.240**	-0.228**	-0.267***	-0.228**	-0.240**	-0.276***
	(0.095)	(0.084)	(0.094)	(0.096)	(0.098)	(0.094)	(0.094)	(0.101
Only Women	-0.029	0.023	-0.032	-0.028	-0.042	-0.021	-0.025	-0.068
	(0.085)	(0.074)	(0.085)	(0.084)	(0.082)	(0.084)	(0.083)	(0.078)
Group Started Amongst…								
Friends								0.239***
								(0.079)
Members of same trade								0.280**
								(0.126)
Neighbours								0.175
								(0.120)
Other								0.141
								(0.116)
Members of another group								0.207*
								(0.117)
President/Committee Decides	0.083	0.134**	0.080	0.084	0.064	0.079	0.077	
	(0.067)	(0.061)	(0.067)	(0.068)	(0.068)	(0.067)	(0.069)	
President/Committee Paid		0.974						
		(0.659)						
*President/Committee Paid *Decides*		-1.029						
		(0.676)						
Survey			-0.052				-0.050	
			(0.087)				(0.087)	
Known				0.006			-0.001	
				(0.059)			(0.059)	
*Survey * Known*					-0.069			
					(0.067)			
Other Conditions						0.030	0.026	
						(0.072)	(0.074)	
Pseudo-R^2	0.56	0.59	0.56	0.56	0.57	0.56	0.57	0.59
Observations	97	97	97	97	97	97	97	97

** p<0.1; ** p<0.05; *** p<0.01. Average marginal effects shown. Robust Standard errors in parentheses*

In column 8, the positive and significant coefficient attached to *Started amongst friends* and *Started amongst members of the same trade* suggests that such ROSCAs were more likely to have experienced enforcement problems than those started amongst family members. A group started amongst friends was 28-35% more likely to experience enforcement problems than one started with family members. A similar result emerges with groups started amongst members of the same trade and, to an extent, ROSCAs started amongst members of another group. This result seems to suggest that where strong social ties exist, at least between the founding members of a group, (and as such there is a higher

level of existing social capital) there is a smaller likelihood of enforcement problems occurring.

The variable *Number of cycles completed* captures how many cycles each ROSCA had successfully completed before the 2004 survey. One would expect the likelihood of enforcement problems occurring to be lower in groups that have completed more cycles due to familiarity and trust between members having been built up over many rounds of previous interactions. The results present some evidence to suggest that duration of existence has a role to play; the variable is often negative and significant at the 1% level: completing a further cycle might reduce the likelihood of enforcement problems occurring by around 1-1.5%. However, this result is not robust to the inclusion of all other controls. An alternative variable to include here might be the duration of existence – some groups had existed for almost 40 years, whilst some were only recently formed in 2004. Results were similar to that of the number of cycles.[26] In section 2.5 below, further inquiry is made with regard to the role of group age, asking whether or not repeated interactions and completing more cycles can help to overcome the odds of experiencing enforcement problems.

The estimations also suggest that those groups that had experienced and survived problems before 2004 were more likely to see a recurrence; it appears common that a group which has experienced enforcement problems already will do so again, adding weight to the picture of instability and riskiness associated with investing one's savings in a ROSCA. However, this also suggests that many groups are resilient to such an occurrence.

Further controls are included for groups whose membership was restricted to the same ethnicity or gender by including three dummies (*Same Ethnicity, Only Men* and *Only Women*; the reference categories are thus groups allowing members of any ethnicity or gender respectively to join). The results suggest that those ROSCAs composed purely of men (23 of 97) seemed to have a lower likelihood of experiencing enforcement problems than mixed groups. Anecdotal evidence from several members of different groups indicated that mixed-gender groups can be more difficult to manage. According to some women interviewed, "men can often take too much space" and adultery issues have marred some groups.[27] Spouses in Benin have also been observed to be secretive about their incomes and

[26] Not shown.
[27] This is why, in order to avoid potential disputes concerning mainly adultery issues, around 40% of groups allowing female membership require husbands' approval for new female members.

have disconnected financial spheres (LeMay-Boucher and Dagnelie, 2014); this provides some indication as to why single gender groups appear to be more sustainable.

Conjecture 9: *Enforcement problems are more likely to occur in larger groups due to higher monitoring costs and a lower level of social capital, on average, between members.*

The discussion of size above highlighted various relationships between the number of members and the likelihood of experiencing enforcement problems. Tables 2.2, and 2.3 show that there is a negative relationship between size and enforcement problems; that is, in larger groups the likelihood of experiencing problems is actually lower. The magnitude of this impact is relatively small compared to others discussed above (i.e. written rules). This result might go some way to corroborating the argument that larger groups are better at dealing with a member who cannot make his or her payment, either because the contribution itself is small, or purely because there are a higher number of others who could potentially step in to help or monitor. Another factor that might explain this result is that social sanctions (notably loss of reputation) are likely to be greater in larger groups in case of default.

2.4.3 Robustness checks

The surveys in Cotonou did not investigate the root cause of each enforcement problem in detail; a group might, of course, be tightly knit, well designed, and effectively screen new members, but there is always the chance that ROSCAs experience problems or collapse due to one or more members falling into financial hardship or facing an unexpected income shock. It was shown above that groups which allocated funds based on members' need were significantly less likely to experience problems, however the majority of groups did not carry out this function and as noted, this variable might represent those members requiring the collective savings at set intervals to cover regular lumpy expenditures. Whilst it is not easy to control for the likelihood of an income shock, information at the individual level allows us to make some attempt to control for the personal circumstances of the membership.

Table 2.4 *Robustness Checks; Dependent Variable =1 if ROSCA experienced enforcement problems between 2004 and 2006*

	1	2	3	4
Membership size	-0.009***	-0.008***	0.008	0.009
	(0.002)	(0.003)	(0.010)	(0.011)
Membership size ^{2}			-0.000*	-0.000*
			(0.000)	(0.000)
No. of Cycles Completed	-0.012***	-0.013***	-0.012***	-0.013***
	(0.004)	(0.005)	(0.004)	(0.004)
President	-0.276***	-0.280***	-0.274***	-0.295***
	(0.085)	(0.099)	(0.091)	(0.098)
Random	-0.020	-0.011	-0.036	-0.038
	(0.076)	(0.082)	(0.077)	(0.083)
Need	-0.502***	-0.663***	-0.610***	-0.658***
	(0.114)	(0.225)	(0.192)	(0.208)
Rules	-0.227***	-0.266***	-0.259***	-0.279***
	(0.070)	(0.078)	(0.075)	(0.081)
Experienced Past Problems	0.762***	0.932***	0.872***	0.940***
	(0.105)	(0.219)	(0.179)	(0.193)
Pot size (CFA 000's)	0.000*	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Monthly Meetings	-0.167**	-0.151*	-0.136*	-0.146*
	(0.067)	(0.080)	(0.074)	(0.080)
Single Ethnicity	0.030	0.046	0.083	0.090
	(0.072)	(0.078)	(0.075)	(0.081)
Only Men	-0.232**	-0.255**	-0.228**	-0.246**
	(0.094)	(0.105)	(0.097)	(0.105)
Only Women	-0.018	-0.035	0.002	0.002
	(0.082)	(0.089)	(0.080)	(0.086)
Interviewee(s) held job for > 2 years	0.011 (0.087)			
Contributions' share of total expenditure	-0.074 (0.068)			
Pseudo-R^2	0.57	0.53	0.57	0.55
Observations	95	90	97	90

** p<0.1; ** p<0.05; *** p<0.01. Average marginal effects shown.*
Robust Standard errors in parentheses

Column 1 of table 2.4 replicates the basic specification of table 2.2, including controls for the personal characteristics of the member(s) of each group observed. The variable *Interviewee(s) held job for > 2 years* is a dummy equal to 1 if the member(s) surveyed had held their current job for at least 24 months (in 2004); *Contributions' share of total expenditure* is the share of an individuals' total expenditure taken up by ROSCA contributions for the average member in each group. The first is included as it might reasonably be expected to capture the stability of an individuals' income and thus the likelihood that they will face an income shock, reducing their ability to make ROSCA contributions. The second acts as a measure of how important contributions are in an individuals' budget. In the face of an income shock, ROSCA contributions might not be a priority to an individual, especially if they take up a large portion of their income.[28] The majority of groups in this study were observed only once (i.e. only one member of each group was interviewed), however 20 of the 97 were observed two or more times. Where there is information on more than one member, an average is taken. The inclusion of these

[28] As a further robustness check, other individual characteristics (age, education level, etc.) were added to the model. Results not shown confirmed that none of these additional variables had any impact on the results presented above.

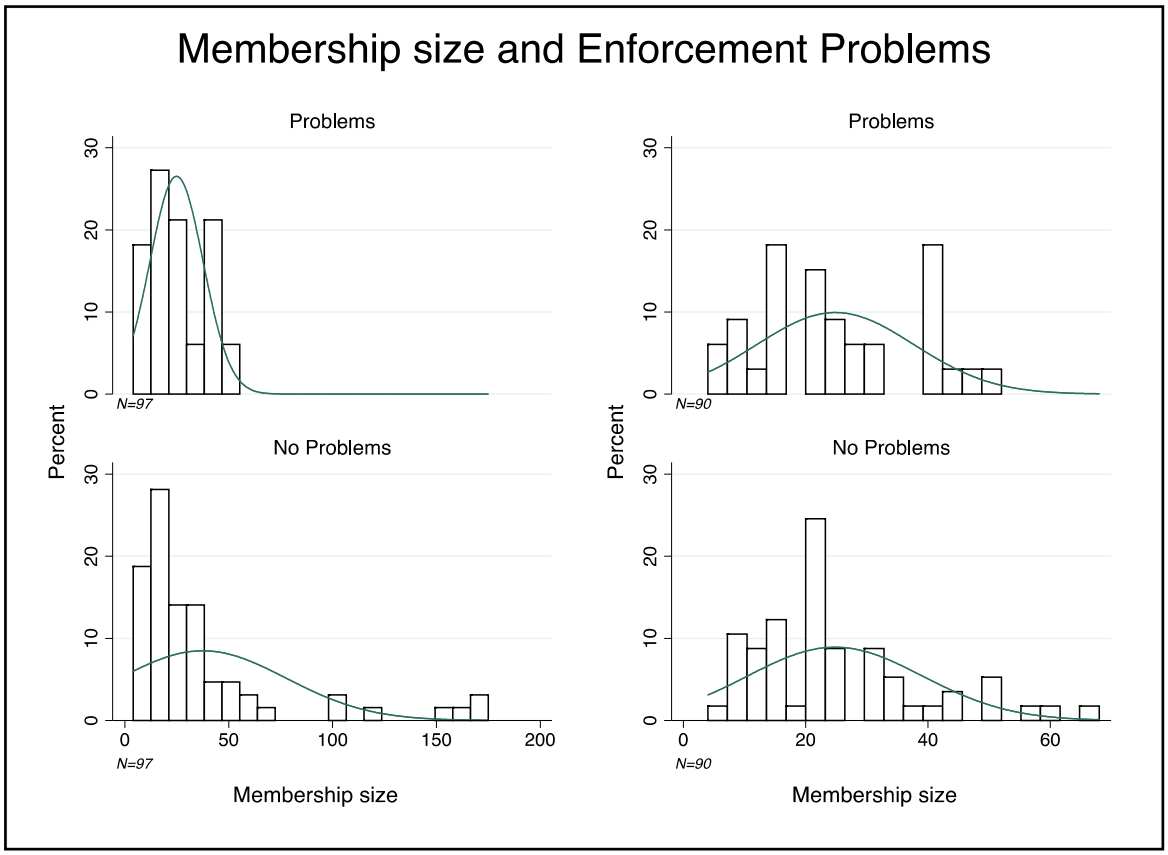
variables is based on the assumption of homogeneity of membership: a feature of ROSCAs emphasised in the literature. They are often formed amongst people from '…the same office' (e.g. Besley *et al*, 1993; Handa and Kirton, 1999), with 'similar occupations and income' (Ambech and Treich, 2007) or 'the same community' (Anderson and Baland, 2002). Table 2.2 highlighted that around 19 per cent of groups in the sample were started by individuals from the same neighbourhood; 43 per cent amongst friends and 14 per cent amongst members of same trade. All of the previously discussed results are robust to the inclusion of these controls, but they do not carry any explanatory power.[29]

Returning to the issue of group size, the left hand panels of figure 2.2 display the distribution of membership sizes according to whether or not the group experienced enforcement problems. Whilst there is quite a large distribution of group sizes, the majority of these are clustered between zero and 75 members; indeed, it appears that there are some clear outlying groups with very large membership sizes. Crucially, none of these groups had experienced problems in the years between our surveys, but we cannot rule out the possibility that these larger groups are causing a bias on the coefficient for membership size. The right hand panels of figure 2.2 show the distribution, having trimmed the dataset to remove the largest groups. These histograms suggest that the distribution of group size between those experiencing enforcement problems and those not is quite similar. Column 2 of table 2.4, replicates the basic specification in column 1 using a trimmed dataset: all of the previous results hold.[30]

[29] These variables were also included in all previous specifications (not shown) but the results remained unchanged. The sample is reduced to 95 observations here due to missing data on two members. A test for equivalence of means between those groups that did and did not experience problems showed no significant differences in these two variables.

[30] The sample was just trimmed in an *ad-hoc* manner by eyeballing the distribution of group sizes.

Figure 2.2 *Membership size and Enforcement Problems*



I also investigate the possibility of nonlinearity between size and the dependent by including the square of membership size (column 3). The results do not provide much evidence of a quadratic relationship. When including the square of size, trimming the dataset to 90 observations has no discernible effect (column 4).[31] It is also possible to take account of outliers in the discrete series by specifying them in natural log form; all of the previously discussed results remain robust.[32]

One further potential limitation of the above analysis is that it is not possible to control for the fact that individuals self-select into groups, leading to a kind of endogenous group formation; most often they are free to choose which they join and in some cases might even have control over the institutional design of the group before commencing making payments. With this in mind, it is perhaps likely that individuals who know that they are reliable and do not plan to default may select a group with, for example, a president, written rules, or one with no history of problems; conversely, any member with ideas of defaulting

[31] These results are robust to different *ad hoc* trimming rules where, for example, a similar (5-10; instead of 7) number of ROSCAs were identified as outliers.
[32] Results not shown but available on request. I prefer not to impose the logarithmic transformation to simply take account of outliers; furthermore, the pseudo-R^2 was larger when specifying these variables in their linear form.

for his or her own profit might choose a group with a less formal structure. However, such concerns are not likely to be of great importance, for a number of reasons. Firstly, 60 % of the interviewees data stated that, given the choice, they preferred to receive the pot towards the end of the cycle, whilst fewer than one in four preferred to receive the pot at the beginning. Dagnelie and LeMay-Boucher (2012), using the related Beninese household sample, show that ROSCA participation is most often used as a commitment device – thus the majority of individuals have some reason to *want* to save and commit to the ROSCA arrangement. The preference for commitment is not correlated with duration of ROSCA membership and thus unlikely to be the effect of some learning process. Secondly, it is unlikely that individuals have perfect knowledge of the reliability of different groups (and the effect of group design on this). As a result, they are more likely to simply join the group most convenient to their own situation, considering the size of contributions, place and regularity of meetings or whether or not they know a current group member. It is, however, difficult to interpret the results here as causal. Without very detailed information on every member of every group, it is not possible to say whether groups experienced enforcement problems due to the composition of the group (i.e. characteristics of the individuals) or due to the incentives created by the group design. Whilst controls for individual characteristics of a representative member are included above as a check, it is unlikely that these capture all information on individuals' circumstances.[33] Thus, the results can only confidently be interpreted as correlations. Future research might wish to improve on the design presented herein by examining the individual characteristics of more, if not all, members of each group over time and inquiring further as to the exact cause of each case of non – payment. Detailed information such as this would assist in shedding light on the exact role that the institutional design and sociocultural makeup of ROSCAs plays in helping to avoid default.

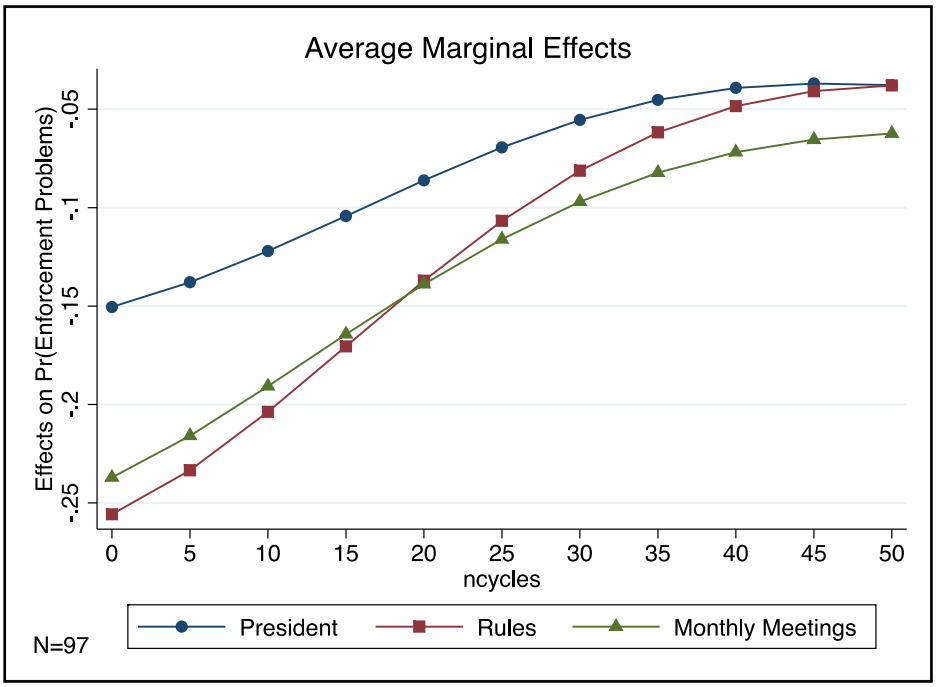
2.5. Further analyses

2.5.1 Given a certain institutional design, can repeated interactions lower the likelihood of experiencing enforcement problems?

[33] Further evidence in this regard comes from Kedir and Ibrahim (2011), who found that the characteristics of a ROSCA (e.g. group size, frequency of draw) were not significantly related to the size of contribution. To the extent that the size of members' contributions reflects the *type* of group, or the intentions of members (i.e. those contributing larger sums to a group might be more motivated to ensure that their savings are protected), then these results suggest that the institutional design of groups is not correlated with the motivations of members to make payments or not.

In all previous results, the average marginal effects were shown. However it is possible to compute marginal effects of covariates of interest at set levels of one of the other independent variables.[34] Of particular interest is how the effects of some of the institutional features might change as ROSCAs complete more cycles (or get "older"). That is, as opposed to asking "*What is the average effect on the likelihood of enforcement problems of having a president (as opposed to a committee)?"* one can ask, *"What is the effect on the likelihood of enforcement problems of having a president (…) for a group that has completed X cycles?"*

Figure 2.3 *Marginal Effects of Institutional Features measured at different stages in the ROSCA life cycle*



The plot in figure 2.3 displays the marginal effects of three different institutional features on the probability of experiencing enforcement problems, measured at 5 cycle intervals. The features chosen are having written rules, ruling structure and monthly meetings. These three features were selected as they were all (i) institutional design features that can be chosen by a ROSCA and (ii) significantly related to the likelihood of enforcement problems occurring in the main regression results. The plot shows that for 'young' groups, having a president is the most important institutional feature of the three (followed by having less frequent meetings and having written rules) for deterring enforcement problems. However interestingly, after having completed around 20 cycles (which is the case for 6 groups in

[34] The Stata commands *margins* and *marginsplot* are used here.

our sample), it appears that having written rules actually acts as a stronger deterrent to enforcement problems than having infrequent meetings.[35] Key here, however, is that the importance of all three features declines as groups get 'older'; the marginal effect of all three declines to between -4 and -8% after around 30 cycles have been completed.[36] This suggests a diminishing importance of institutional features as groups get older, and support for the theory that repeated interactions amongst members can strengthen the bonds of trust and reciprocity within the group.

Figure 2.4 *Marginal effects of interactions of Institutional Features measured at different stages in the ROSCA life cycle*

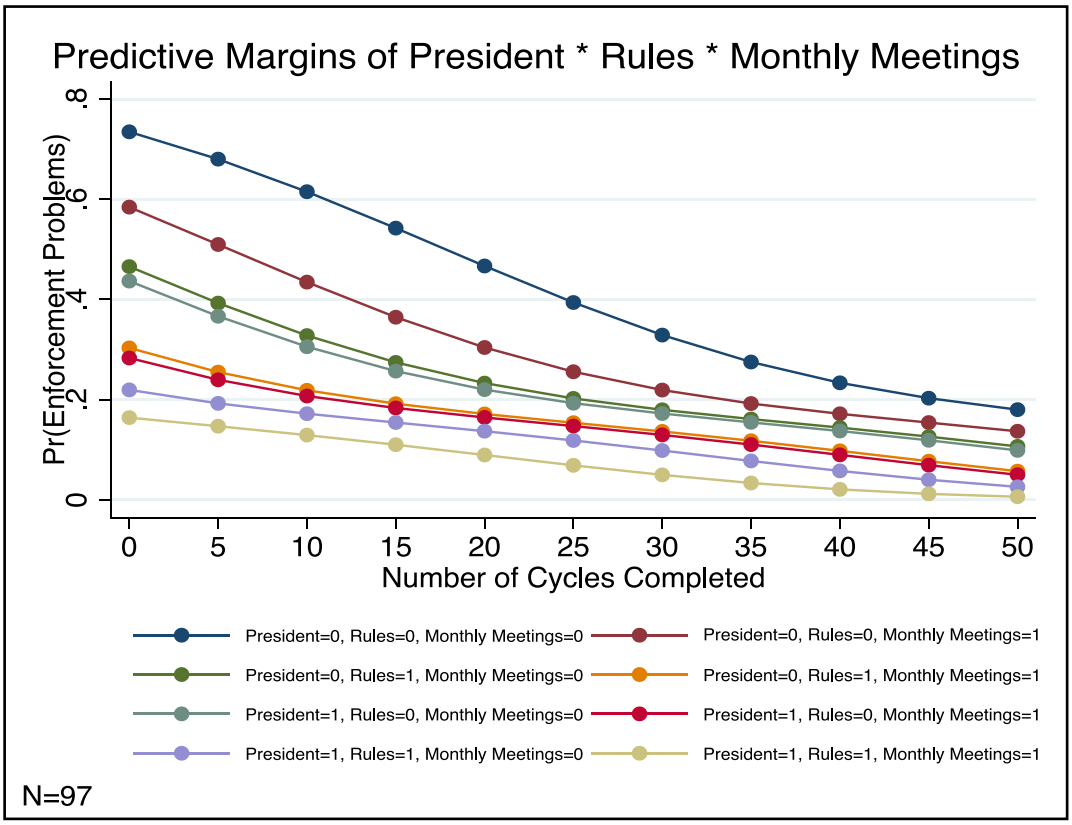


Figure 2.4 plots the predictive margins of interaction terms between the three variables and as such allows for an examination of those combinations of features that might be more or less likely to lead to enforcement problems. It is clear that having a committee (i.e. *President* = 0), no written rules and meeting more than once a month left groups most susceptible to problems; the marginal effects suggest that, at between 0 and 5 completed cycles, these groups were around 70-75% more likely to experience enforcement problems

[35] The average cycle length lasts for 13 months. The variable 'number of cycles' varies between less than 1 and over 130, with a mean of 8.6.
[36] The marginal effects of 'monthly meetings' remains significant up to 50 cycles. The marginal effect of 'written rules' is significant up to about 35 cycles. The marginal effect of 'president' is only significant up to around 20 cycles.

than one with any other combination of the three features. However, again it appears that groups facing this initial disadvantage can, to an extent, overcome it if they survive the early cycles. After having completed 20 cycles, the marginal effect suggests that these groups were 47% more likely to have experienced enforcement problems, representing an improvement in their chances of not doing so, compared to some other structure, of over a third. Of course, it is true that for all combinations of the aforementioned features, increasing the duration of the group decreases the likelihood of experiencing enforcement problems (all of the plots slope downwards). But at, say, 50 completed cycles, the importance of institutional features appears greatly reduced (as does the difference between having different combinations of features), with even those groups facing the greatest initial disadvantage able to enjoy much greater odds of not experiencing enforcement problems.

2.5.2 What influences the likelihood that a group, having experienced enforcement problems, will collapse?

The above analysis has considered the influence of various factors on the likelihood of experiencing enforcement problems. Table 2.5 presents results from a test of equivalence of means between groups that survived (10) and those that collapsed (23) following an enforcement problem.[37] Immediately apparent is that there are few significant differences – only having written rules appears to be significantly higher amongst those groups that survived. On average, those groups that survived were slightly larger (29 v 23), had completed more cycles (3.5 v 3), and were less likely to have a president. The remainder of the variables, whilst not displaying statistically significant differences, seem to paint a similar picture as above. Groups that survived enforcement problems had, on average, more members, completed more cycles, and written rules. This might be tentatively interpreted as a sign that the same characteristics that help to minimise the likelihood of enforcement problems occurring also help groups to survive in the face of such problems, should they nevertheless occur. Only half of those that collapsed had survived enforcement problems in the past, whilst seven out of the ten surviving had done so, suggesting that some groups are used to non-payment and have devised means with which to deal with it.

[37] Both Probit and Heckman analyses were also attempted for this subsample, but neither yielded any meaningful results due to the small sample size.

Table 2.5. *Equivalence of means; Groups that experienced enforcement problems – survived v collapsed*

Variables	Collapsed		Survived		Difference
	Mean	*S.E.*	*Mean*	*S.E.*	
Membership size	23.044	(2.63)	29	(4.137)	-5.957
Number of complete cycles	3.040	(0.587)	3.528	(0.512)	-0.489
Duration of Existence (months)	38.91	(11.97)	57.4	(13.42)	0.369
Only President	0.348	(0.102)	0.100	(0.100)	0.248
President/Committee paid	0.174	(0.081)	0	(0.000)	0.174
Random	0.609	(0.104)	0.700	(0.153)	-0.091
Need	0.087	(0.060)	0	(0.000)	0.087
Other Fixed	0.304	(0.098)	0.300	(0.153)	0.981
Written Rules	0.348	(0.102)	0.700	(0.153)	-0.352*
Monthly meetings	0.348	(0.102)	0.300	(0.153)	0.048
More severe sanctions on delinquent member	0.652	(0.102)	0.900	(0.100)	-0.248
Pot size (1000's of CFAF)	98.165	(21.211)	122.3	(31.664)	-24.135
Single ethnicity	0.304	(0.098)	0.200	(0.133)	0.104
Only men	0.304	(0.098)	0.100	(0.100)	0.204
Only woman	0.174	(0.081)	0.100	(0.100)	0.074
Survived Past Problems	0.391	(0.104)	0.700	(0.153)	-0.309
Started amongst…					
Family	0.043	(0.043)	0.010	(0.100)	-0.057
Friends	0.522	(0.107)	0.800	(0.133)	-0.278
Members of same trade	0.174	(0.081)	0	(0.000)	0.174
Neighbours	0.174	(0.081)	0.100	(0.100)	0.074
Other	0.044	(0.044)	0	(0.000)	0.044
Members of anther group	0.044	(0.044)	0	(0.000)	0.044
President/Committee decides	0.652	(0.102)	0.700	(0.153)	-0.048
Survey on New members	0.739	(0.094)	0.700	(0.153)	0.039
New members must be known	0.522	(0.107)	0.800	(0.133)	-0.278
Other conditions	0.783	(0.088)	(0.900)	(0.100)	-0.118
N = 97	*N=23*		*N=10*		

2.6 Conclusion

This chapter has presented the first evidence on the sustainability of ROSCAs over time. The empirical investigation tests some of the intuitions outlined in the related literature. Knowledge of the norms in ROSCAs, along with how the groups survive over time, is of high importance if we wish to fully understand how these groups function and advise policymakers, MFIs, NGOs or formal banks on the future of savings and credit mobilisation in developing countries.

Whilst ROSCAs are often cited for their remarkable stability in the face of clear incentives for members to default, the evidence presented here suggests that this is not necessarily always the case: one in three groups in the sample from urban Benin experienced enforcement problems within a two-year period; around a quarter of the sample collapsed in the same time. The empirical investigation sought to highlight the institutional design of those groups that were more or less likely to have experienced enforcement problems and a number of interesting insights emerge. The results suggest that those groups run by a president, those with written rules and those that fixed the allocation of the fund according to member's need were less likely to have experienced enforcement problems between

2004 and 2006. This latter result – on the order of pot reception- suggests that pragmatism may be key: responding to the financial needs of members, rather than allocating the pot according to a random draw or some other fixed process, appears to foster stability within ROSCAs. The results fail to provide support for the theory outlined in Anderson *et al.* (2009), which suggested that random allocation ROSCAs have more severe enforcement problems than fixed ROSCAs. It might also be seen a reflection of the importance of the commitment motive for membership, as outlined in Dagnelie and LeMay-Boucher (2012). Furthermore, evidence is presented that groups meeting less frequently, or those that were larger, were less likely to experience enforcement problems, corroborating some of the predictions of the model in Besley *et al.* (1993). However, I find only weak evidence that groups with a larger pot were more likely to have sustainability problems. With regard to the importance of social capital and how this can promote cohesion amongst group members, results suggest that groups started with family members are less likely to have experienced problems compared to those started with friends or colleagues. This corroborates the notion that groups with stronger existing ties will be closer-knit and therefore less likely to experience enforcement problems. The results here are also comparable to those in, for example, Karlan (2007), which investigated the impact of social connections on repayment rates in village banks in Peru, finding that those groups with higher levels of social capital performed better. However, little evidence is found that formal screening arrangements reduce the likelihood of enforcement problems occurring. An analysis of the importance of institutional design over time highlights that groups facing an initial disadvantage can, by successfully completing more and more cycles, lower their likelihood of facing default. Repeated interactions can help to reinforce group stability.

Chapter 3: Exploring Regional and Gender Disparities in Beninese Primary School Attendance: A multilevel approach

3.1 Introduction

Benin has seen almost unparalleled improvements in primary school attendance since 1990, yet remains virtually ignored in the literature surveying education outcomes in developing countries. Gross Enrolment Rates (GERs), defined as the number of students attending school as a percentage of the school age population, have soared from around 50% in 1990 to well over 100% in 2012.[38] The *Net* Enrolment Rate (NER), defined as the percentage of *school age* students attending school, stood at 95% in 2012 (UIS, 2105) whilst the average in Sub-Saharan Africa (SSA) was just 77%. However, despite these impressive increases in enrolment, severe geographical and gender disparities remain present in Benin's primary school attendance rates. According to the most recent DHS data, there are still communes where less than one third of children attend.[39] Of course, the reasons why attendance is likely to be lower in rural areas, or for girls, are well understood: lower perceived benefits coupled with higher costs (or opportunity costs) of attendance mean that parents in rural areas, or parents of daughters, are often less likely to send their offspring to school.

This study complements individual- and household-level data from the Beninese Demographic and Health Surveys (DHS) with detailed commune-level schooling data from the Beninese Institute for Statistics (INSAE) in order to assess the importance of demand- (i.e. on the part of the student, or their family) and supply-side factors (i.e. availability of school facilities) on school attendance rates. Results from a logistic regression model suggest that household wealth, religion, parental education and the supply of schools all predict the likelihood of a child attending school. Results also indicate that as average distance to school increased, the likelihood that boys who work in the field would attend school decreased, to a greater extent than those who did not work. This echoes the findings of, for example, Lincove (2012) or Huisman and Smits (2009) and also presents evidence that distance to school acts as a useful proxy for the opportunity cost of attending, which is greater when time travelling to school replaces potentially-income generating work.

[38] Impressive GERs, however, cannot be taken at face value. By definition, a value greater than 100 might point to a system playing catch-up, or a large number of students entering late or repeating grades. Both are likely to apply in the Beninese context.
[39] Benin is subdivided into 12 *departments* and 77 *communes*.

Whilst much of the existing research on primary school attendance (e.g. Huisman and Smits, 2009) has acknowledged that factors at the community, district or national level play an important role in explaining school attendance, few have sought to explicitly model this econometrically. As a result, estimations fail to account for cluster-level interdependence. This study also employs a multilevel logistic model, which accounts for unobserved heterogeneity between higher level clusters (i.e. at the household and commune level), in order to assess the level at which most variation in school attendance rates is seen. A three-level *random intercepts* model shows that there are a large number of communes with a significantly lower than average primary attendance rate. However, only around 11% of this variation is attributable to factors at the commune level. After controlling for the number of schools and average distance to school, less than 5% of the variation in attendance is due to commune-level factors, suggesting that the majority of the between-commune variation in attendance rates is attributable to factors at the household level. These results are important from a policy context: it is useful for policymakers to be able to pinpoint those levels at which most variation in attendance rates exists and as such those areas where interventions might be most effective (i.e. at the household, community, or commune level). Following Delprato and Sabetes (2015), a *random slopes* model is also estimated, where the wealth coefficient is allowed to vary between communes. This pinpoints those communes where attendance is below average, but the effect of household wealth on school attendance is above average. Policymakers could, for example, use such results in order to target those areas where interventions that raised household wealth (or lowered the cost of schooling) might be most effective in increasing school attendance rates.

This study also investigates the inherent problems with measurement of school enrolment or attendance; where public officials such as school teachers or principals have an incentive (due to top-down funding replacing school fees, as was the case in Benin ten years ago) to report higher enrolment rates, then official statistics might inflate the true number of attendees. Indeed, evidence is presented that enrolment figures from the Beninese DHS are somewhat lower than those from UIS (UNESCO Institute for Statistics), or INSAE.

The broad contributions to the related literature are therefore threefold: Firstly, evidence is presented on the determinants of primary school attendance for Benin, a context that has not previously been considered in the empirical literature surrounding primary schooling.

Secondly, the use of a multilevel model helps to provide additional insights on regional disparities that many similar studies, from other countries, neglect to consider. Finally, the comparison between official statistics and household survey data provides further detailed evidence of the problems with using official statistics, noted in a number of other studies.

The rest of this chapter is structured as follows. Section 3.2 provides a discussion regarding the disparities in enrolment statistics from different sources, before considering the Beninese context in detail. Section 3.3 discusses the theoretical predictions and empirical results surrounding school attendance in developing countries, with an emphasis on not only the economic rationale, but also the sociocultural factors that might dictate whether or not a child is sent to school. In particular, results from similar studies in Sub Saharan Africa (SSA) are surveyed. Section 3.4 presents the variables and methodology chosen for this chapter. Results for the single-level logistic model are presented in section 3.5, whilst the multilevel strategy is outlined and presented in section 3.6. Section 3.7 concludes.

3.2 Measuring Attendance: The Beninese Context

3.2.1 Data considerations

The various sources reporting enrolment or attendance statistics for the period in question in Benin appear to tell a somewhat different story. Table 3.1 illustrates that official enrolment statistics, from UIS or INSAE are consistently higher than those from the DHS.

Table 3.1 *Differences in enrolment statistics by source.*

Source	*Indicator*			
	NER (%) 2006	NER (%) 2012	GER (%) 2006	GER (%) 2012
DHS	57.01	71.05	86.16	96.41
UIS	82.58	94.86	98.79	122.77
INSAE[40]	-	-	99.59	-

However, there are a number of reasons why statistics from the DHS might not only differ from those supplied by the government (INSAE, UIS[41]), but might also be a more reliable and useful indicator of school attendance. Over-reporting on the part of public officials can lead to upward bias in enrolment statistics (Glewwe and Kremer, 2006; Sandefur and

[40] INSAE does not report NER and the data for GER 2012 is missing.
[41] Whilst not identical, the UIS and INSAE statistics track each other very closely.

Glassman, 2015). Sandefur and Glassman (2015) investigated such a bias toward over-reporting enrolment statistics in a panel of 46 surveys in 21 African countries, finding this bias to be prevalent in cases where low-level public servants (in this case, teachers or school principals) had incentives to misreport official statistics. This was particularly true in countries such as Kenya or Rwanda, where pupil fees had been replaced by top-down per pupil grants – exactly the case in Benin during the period in question here. These authors, and also *fhi360* (2013), point out that part of the discrepancy between *enrolment*, which is reported in official statistics and *attendance*, which is measured by the DHS, might also arise from children that enrol in school but rarely attend. Indeed, the fact that enrolment is now free in Benin might lead some to enrol with little intention of ever attending. Thus, due to the potential question marks raised over the validity of official statistics, this study will rely on the DHS data where possible, as these are widely used, understood and accepted as representative. The respondents are unlikely to answer strategically, compared to public officials that might have incentives to over-report the number of children enrolled in order to maximise funding for their own schools. Moreover, an indicator of *attendance* is perhaps a more useful indicator than *enrolment* in the sense that it captures the number of children who actually attend school, rather than those merely registered to attend. Unfortunately, however, the DHS surveys do not investigate the frequency of attendance at school, asking only whether a child attended or not.

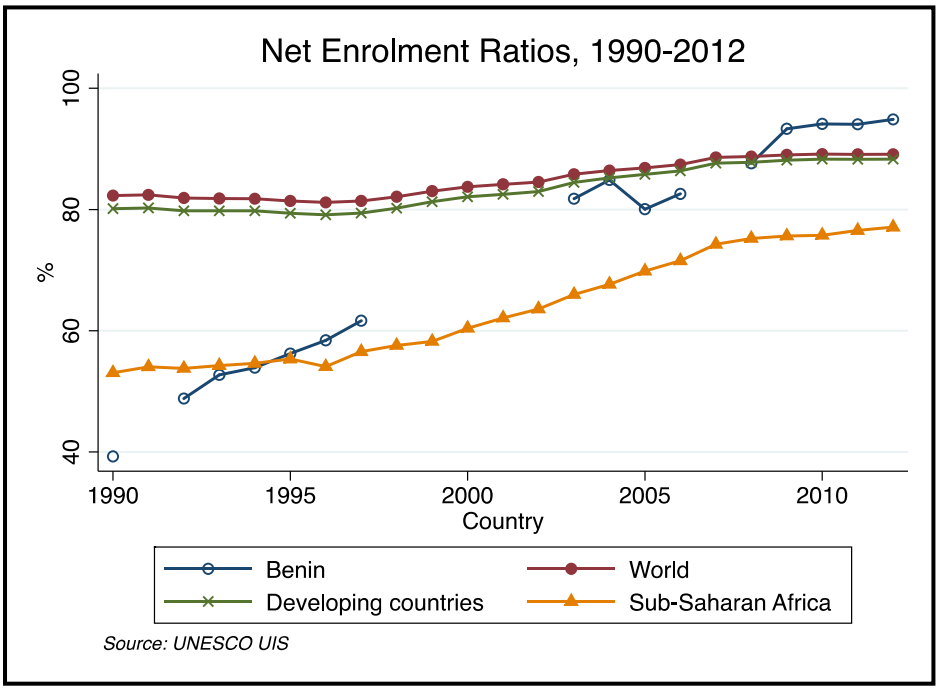
3.2.2 The Beninese Context

Benin provides an intriguing case study in education and development. Along with an economic crisis that forced the closure of teacher training colleges and large scale cuts to the civil service in the 1980's, the socialist regime's failed attempts at reform (see Allen, 1989) left Benin ranking amongst the worst performing countries in the world with regards GERs and gender parity; fewer than 50 per cent of all children and less than one in three girls were attending school in 1990; the ODI (2011:4) describes the education system at this time as "deeply dysfunctional and inequitable". However, the democratically elected government prioritised education in 1990 and as a result of systemic reform, Benin has seen almost unparalleled (at least in SSA) progress in terms of enrolment rates.[42] Between 1990 and 2010, average adult years of education (Barro and Lee, 2013) increased from 2.13 to 4.35 (only Mali saw a greater increase in the same period), gross and net enrollment soared and the gender gap was virtually eliminated in many regions. In 2006, primary education

[42] For a more detailed account of the backdrop to the 1990 reforms, see ODI (2011)

was made free for all. Figure 3.1 plots Benin's progress in net enrolment rates against the average for SSA, developing countries and the world – despite missing data for many years, the trend is clear: NERs rose from around 40% (52% male, 27% female) in 1990 to 94% (99% male, 88% female) in 2012.[43]

Figure 3.1 *Net Enrolment Ratios 1990-2012*

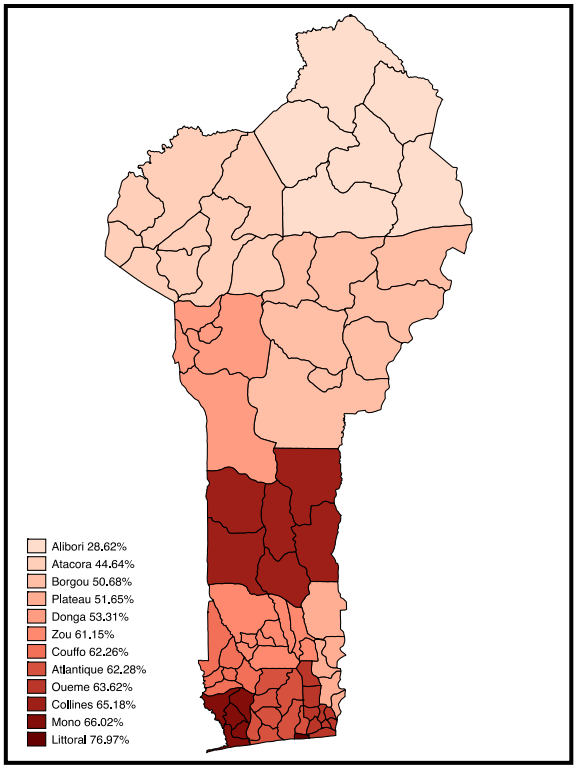


Gross enrolment rose from 51% to 123% in the same period. Progress in increasing enrolment of girls in school has been particularly impressive: the Gender Parity Index (GPI) rose from 0.50 in 1990 to 0.89 in 2012 – only Guinea matched this progress in SSA.[44]

Yet these figures do not tell the whole story in Benin, as large regional disparities persist at both the department and commune level. Figure 3.2 shows the primary attendance (male and female combined) by administrative department in 2006. The data in this case comes from the 2006 wave of the DHS. Given the discussion above regarding statistics from the DHS, these numbers might best be defined as *Net Attendance Rates*; i.e. the percentage of primary school age children whose parents reported that they were attending primary school in 2006.

[43] I rely here on the UIS data as it allows for comparisons with world and regional averages
[44] The GPI is calculated as female gross enrolment divided by male gross enrolment. Thus this statistic equals 1 when gender parity is achieved.

Figure 3.2 *Net Primary Attendance by department, 2006.*

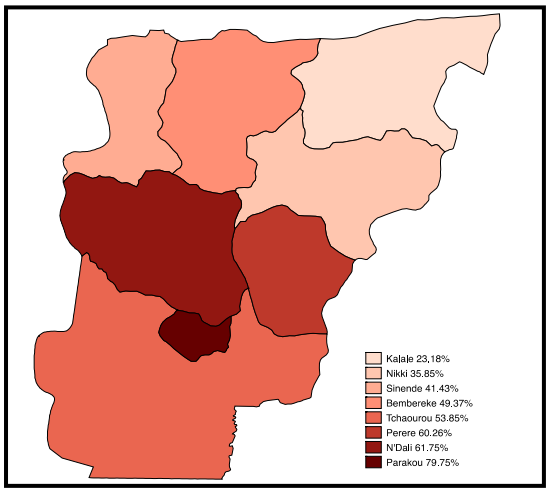


Source: Author's own calculations from the 2006 DHS. (NB. Mean = 57%)

In 2006, some regions were still struggling with net attendance rates of as little as 29% (department average), whilst others saw rates of over 75%. The national average was 57%.[45] It is more interesting still to examine the differences at the communal level. Take the department of Borgou in figure 3.3, for example. Net attendance rates in 2006 ranged from just 23% (Kalale) to almost 80% (Parakou). Whilst the Beninese government has continued to prioritise access to education, regional disparities such as those outlined above persist. Indeed, by 2011-12 many communes had net attendance close to 90%, but some still lagged behind in the 20 – 30% range.[46]

[45] NB. This number differs somewhat to that reported in figure 1, due to reasons outlined in section 4 below.
[46] Unfortunately, INSAE has not made available its school supply statistics for 2011-12, so the main analysis here focuses on the 2005-06 round of the DHS, where complimentary statistics from INSAE are available.

Figure 3.3 *Net Enrollment Rates by commune, Borgou department, 2006*



Source: *Author's own calculations from the 2006 DHS. (NB. Mean = 50.68%)*

Whilst the NER (or in this case, the net *attendance* rate) represents a significantly more useful indicator than *gross* enrolment, it still does not tell the whole story; the official definition (UNESCO, 2012) is

"The number of children of official primary school age who are enrolled in primary education as a percentage of the total children of the official school age population."

It therefore pays no attention to what grade a child is in; an 11 year-old, who is, strictly speaking, of school age, having just entered the first grade is counted as enrolled – yet they have arrived in primary school some 5 years late. Whilst the empirical analysis here does not specifically consider on-time-enrolment, a look at the age distribution of those attending primary school is nonetheless interesting. Figure 3.4 uses data from two waves of the DHS in order to highlight changes in the age distribution of primary enrolment in Benin between the 2005-06 and the 2011-12 school year. In 2011-12, the mode age of primary school attendees was 9 years old, at which over 75 percent of children were attending primary school – this compares to 10 years old in 2005-06.

Indeed, the ages containing the highest percentage of children in school are 6,7,8,9,10 and 11 – pertaining to the six official years of school. In 2005-06, the ages with the highest percentages of children in school were 8,9,10,11, 12 and 13. Given that official school age in Benin is 6 – 11 years old, this might reasonably be taken as indication that by 2011-12, more children were attending and completing school earlier, if not still strictly on time. Some of the biggest improvements come when looking to older children: the earlier survey

showed that, for example, around 17% of sixteen year olds were still in primary school. By 2011 that number had fallen to just 7%.

Figure 3.4 *Primary attendance age distribution: 2005-06 vs. 2011-12*

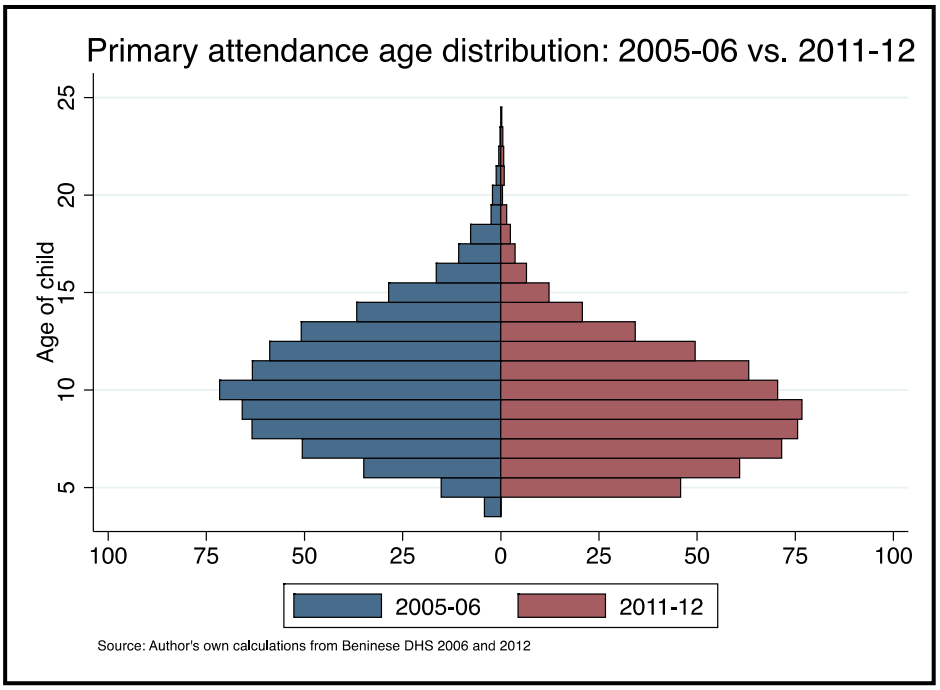
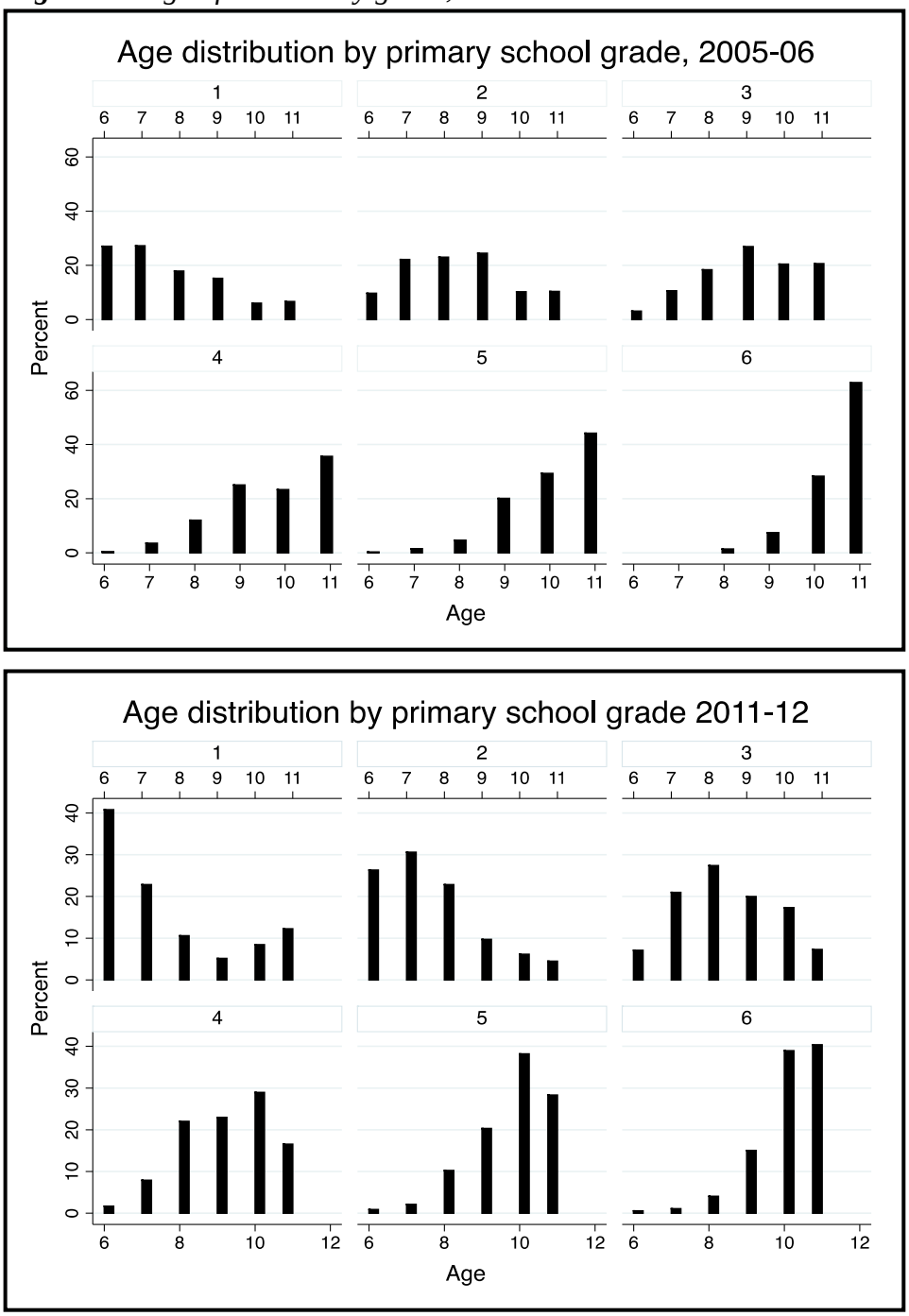


Figure 3.5 goes further, comparing the age distribution of school-age students in each primary grade, in 2005-06 and 2011-12. There is a clear trend toward earlier, or *on time* enrolment of students: The 'official age' for grades 1 to 6 is 6 to 11 years old. The 2005-06 data shows that many students were either entering school late, or playing catch up by repeating grades, with the result that the mode age for each grade was often higher than it should have been. By 2011-12, dramatic improvements have been seen in this regard: with the exception of the 4th grade, the most common age of children in each grade was as expected. Whilst figure 3.5 restricts the sample to those of school age, the conclusions (with regard to the mode age in each grade) are unaltered when accounting for older students, though the spread of the distribution is somewhat wider. [47]

[47] Not shown.

Figure 3.5 *Age of children by grade, 2005-06 and 2011-12.*



Source: Author's own calculations from the 2006 & 2012 Beninese DHS

3.3. Theoretical Predictions and Empirical Evidence

3.3.1 Demand Side

Becker (1975:45) argues that "The most important single determinant of the amount invested in human capital may well be its profitability or rate of return." In other words it makes sense to invest in human capital, or in this case send a child to school, if the expected benefits outweigh the costs incurred.[48] Costs, of course, are measured in both direct and indirect terms – the former constitute items such as school fees, books, transport or uniforms, and the latter the opportunity cost of attending (Becker, 1975). The opportunity cost of attending school is often higher in developing countries than elsewhere: children are expected to work in order to contribute to total household income, particularly so in rural areas. Often times, children enrolled at the start of the school year are pulled out of school and required to help with the harvest in autumn (Colclough *et al*., 2000). Children might make a direct contribution to household income, whereby the child works on the farm or in the marketplace, or an indirect contribution, where their help around the home or family business frees up the time of adults to earn more money (Colclough *et al*., 2000). Empirical work often proxies cost of schooling with a measure of distance to school; time spent travelling to school could be used, for example, to work on the family farm, or help with household chores. Studies such as Lincove (2012) or Huisman and Smits (2009) found distance to school to be inversely related with the likelihood of attendance; Delprato and Sabates (2015) however found no effects on the likelihood of late entry in Nigeria.[49]

Turning to the benefits of attending primary school, these might not be immediately clear to parents or students in a developing country context. If there is a lack of job market opportunities in an area, then there will be a limited expected return to education. Even if opportunities were to arise in the future, parents might not reasonably be able to foresee this happening. If a child is expected to, for example, work on a family plot of land, then numeracy and literacy skills might well be of limited value, at least compared to the physical strength that he or she could have been building, which may prove of more use for his or her future work. More generally, the majority of jobs in an area might not require a

[48] This is in a household production function framework, where parents are deemed to make investment decisions on the part of all household members

[49] Often times it is not possible to have a precise measure of distance between house and school, so approximations based on population and area of a state must be made (e.g. Huisman and Smits, 2009). When even this data is missing, a simple rural or urban dummy might be included in estimations and, within reason, pick up some of the same effect.

formal education, or the education offered at schools may be deemed inappropriate for the predominant type of employment in the area. Similarly, Colclough *et al*. (2000) highlight that in contexts where gender discrimination exists and the gender balance in labour markets is skewed in favour of males, that the benefits of education will be lower for girls. Thus, even with equitable access to schooling, there may still be significant challenges to convince parents of the benefits of sending their daughters to school.

Weighing up the costs and benefits of sending a child to school requires a full knowledge of future benefits. In developing countries, it is by no means guaranteed that parents will be able to accurately measure or estimate such benefits. If information on job market opportunities is unavailable, education is deemed unnecessary for rural farm labour, or if families live in a community where very few adults are educated, then parents (especially those who have not attended school themselves) are likely to undervalue the benefits. In a context where either the future benefits of education are unknown, or where parents display time inconsistent preferences (i.e. hyperbolic discounters who undervalue future benefits), the benefits of education will be undervalued and it is less likely that parents will send their children to school. The costs, whether direct or indirect, are more easily observable.

A formal approach to estimating investment in human capital is outlined in Glewwe and Kremer (2006:965), who consider that each household will maximise a lifetime utility function, and that years of schooling for children will be arguments therein, i.e.

$$S = f(\mathbf{Q},\mathbf{C},\mathbf{H},\mathbf{P}) \qquad [3]$$

Where S is years of schooling, **Q** is a vector of school or teacher characteristics, **C** is a vector of child characteristics, such as innate ability, **H** is a vector of household characteristics and **P** represents costs or prices related to schooling.

The authors note some caveats to any empirical work that attempts to estimate this equation using retrospective data: it is inherently difficult to observe innate ability, teachers' motivation, parents' ability or willingness to assist with homework etc. which will lead to biased Ordinary Least Squares (OLS) regression estimates in the absence of suitable instruments.[50] Yet it is important to note that whilst this conceptual framework is very useful, it is precisely these shortcomings that have dictated the focus of recent relevant

[50] For a fuller discussion, see Pp.968 of the aforementioned article.

empirical studies. Many do not try to explain variation in years of schooling, instead focusing simply on *enrolment* or *attendance rates*. In terms of empirical estimation, there are some important differences. Consider the unobservable of innate ability: A less capable or unfocused child might not complete a full course of schooling (or at least complete 4 years, which Glewwe and Kremer (2006:951) deem the "most appropriate for assessing whether universal primary education has been achieved"). However, he or she may still (i) attend school but have to repeat grades or eventually fail, or (ii) be enrolled at school, but not attend with any degree of regularity. The key point is that any regression where enrolment or attendance rates are the dependent variable will not suffer from the bias outlined above. Enrolment ratios or attendance rates merely capture the number of children on a school's register – whether or not they persist or succeed at school is, by definition, not considered.

Costs and benefits must also be weighed in terms of household wealth, income or expenditures. If the costs of education are small in relation to either of these measures, then it is more likely that parents can afford to send their children to school. It is commonplace for household surveys to stratify households into wealth quintiles (the DHS, for example, does this via a principal components analysis); empirical results often find that the likelihood of attending school increases from the lowest quintile to the highest (see, for example Kazeem *et al*., 2010, Huisman and Smits, 2009), although this is not necessarily always the case. Lincove (2012), for example, found no direct effect of household wealth on school attendance in Uganda, but did find that the effects of other explanatory variables varied in magnitude according to wealth quintile. Delprato and Sabates' (2015) multilevel analysis of late entry to schooling in Nigeria highlighted that community- or state-level wealth effects were greater than those at the household level.

The economic rationale for sending a child to school is, however, just one side of the story: there are various sociocultural differences with respect to religion, caste, tradition or tribe that might interlink with economic decisions to dictate the norms followed by parents with regard to education. Indeed, it is probable that these factors might be driven more by norms at the community or state level, than at the individual household level (Delprato and Sabates, 2015). In societies where patriarchal norms persist, parents may place a higher value on the education of boys over girls. This often stems from the lack of social security or pension system, meaning that male children are expected to provide for their parents in old age; concurrently, it is common that daughters join their husband's family at marriage

(Colclough *et al*., 2000; Huisman and Smits, 2009), and thus their own parents will realise no financial reward to their education. However, the findings of Eloundou-Enyegye and Calvès (2006) provide a challenge to this traditional viewpoint: their study suggests that in some contexts, married women often remitted to their own families (e.g. in Cameroon) and that their capacity to do so actually increased, the more educated they were. Thus a potential paradox exists – parents may be unwilling to educate their more educated daughters in the first place, but more educated daughters might actually remit more money. Their study also notes the significant control that women in West Africa, and Benin in particular, have over their own earnings, a finding corroborated by field surveys cited in LeMay-Boucher and Dagnelie (2014), which confirm the existence of disconnected financial spheres between husband and wife in Benin. A further disadvantage for girls is if the nearest school is quite far away; it might be that they are not allowed to attend until they are slightly older, due to the perceived dangers of walking alone, or the physical effort of walking a long distance.[51]

Religion also plays a significant role in the likelihood of children, specifically girls, being sent to school. Csapo (1981) cites the distrust of Western education by Muslims or the more traditional Islamic views on the education of women, as outlined in the Qur'an, as potential barriers to education for girls in Nigeria. Lincove (2015) found that Muslim children in Nigeria were, on average, 23% less likely to attend school – but the effect was more than double for girls (31%) than for boys (15%), a finding echoed by Kazeem *et al.* (2010) who found the same result for Nigeria, although the order of magnitude is dramatically larger – their regressions suggest that Muslim children were five times less likely to attend school than Christians. Lincove (2012), however, found no effect of being a Muslim on school attendance in Uganda and Buchmann (2000)'s regression results found that Muslim children were no less likely to attend school than their Christian counterparts in Kenya. So, whilst theory predicts that different religions or traditions might place different importance on schooling, the empirical evidence for SSA is mixed. It may well be the case that in many countries, parents of all religion are increasingly willing to educate both their sons and daughters.

The education level of parents is also an important consideration: it is highly probable that if parents have attended school themselves, and benefitted from the education received, that they would be more likely to send their own children to school. In terms of the economic rationale, this might allow the parent to better appreciate and estimate the benefits of

[51] See Colclough *et al*. (2000) for a thorough discussion of the barriers that face many girls in SSA.

education for their own child; this may be particularly true for girls if their mother has attained a certain level of education. Lincove (2015), for example, found that mothers' years of education was a significant determinant of school attendance for both girls and boys in Nigeria and Uganda. Huisman and Smits (2009) also uncovered positive effects of both parents' education level of school attendance; the effect of mothers having at least some primary education was stronger for girls' than boys' likelihood of being in school.[52]

Various household or family level factors such as household size, birth order, whether or not a child is adopted, etc. might also be considered as explanatory factors with regards a child's likelihood of attending school. In terms of economic rationale, a larger number of children increases competition for limited household resources (Lincove, 2015). As a result, it might be that children with more siblings cannot attend school as they are required to complete more tasks at home in order to contribute to household income and the direct costs of sending many children to school are obviously higher. At the same time however, a larger number of children might well increase the likelihood that they can attend school as there are more hands to work and contribute to overall household income. Thus the direction of this effect, if it exists, is unclear. Colclough *et al.* (2000) presents evidence that in Ethiopia, the average number of children in a household was higher for school attendees than for dropouts; Glick and Sahn (2006) found that as the number of children in a household increased, there was no impact on the likelihood that a child would attend a public primary school in Madagascar, although it did decrease the likelihood of a child attending private school. Similarly, Lincove (2015) found no effect of the number of children in a household in Uganda and only marginally significant effects for girls in Nigeria. Where the number of children is large, it might be that only older (or younger) children are allowed to attend school – as a result, a number of studies have considered the effects of birth order. Huisman and Smits' (2009) regressions for a sample of 30 developing countries indicate that later-born girls were more likely to attend school than first-born girls and Kazeem *et al.* (2010) found evidence that having at least one older brother or more than two older sisters increased the odds of attending school in Nigeria. Chernichovsky's results (1985) suggest that an increase in the number of children aged 7 – 14 increased the likelihood that a child was in school in Botswana.

[52] Parental level of education is a common explanatory factor of the likelihood of being in school. Other studies finding positive effects of said are, for example, Buchmann (2000), Deininger (2003), Delprato and Sabates (2015), Glick and Sahn (2006), Kazeem *et al*. (2010) and Lavy (1996).

3.3.2 Supply Side

The above discussion has highlighted that numerous factors interlink to dictate whether or not a child will be sent to school by his or her parents. However, to consider only factors on the demand side ignores many of the considerations that a parent might take into account. Including estimates of the supply of schooling, and quality of that supply, is crucial to avoid omitted variable bias in the estimates obtained in empirical analyses. Often, however, institutional data for the total number of schools, pupil teacher ratios (PTR) etc., is only available at the national or regional level, making consolidation with survey data at the community or village very level difficult. As a result, many studies fail to accurately account for the supply of schooling.[53] A measure of distance to school might be considered as a proxy for the supply of schooling. School quality is often measured by the PTR, although this variable is clearly endogenous in regressions where the dependent variable is school attendance rates. Huisman and Smits (2009) employ an estimate of the number of teachers per 1000 children ('Teacher Child ratio'), finding a positive and significant effect on school enrolment. This avoids the endogeneity problem, as the denominator is total number of children of school age in the population, rather than total number of students.[54] Other variables that have been used to represent school quality include the percentage of teachers who are qualified (e.g. Lavy, 1996; Huisman and Smits, 2009), percentage of teachers who are female (often considered key in explaining girls' participation in education e.g. Glick, 2008; Huisman and Smits, 2009) the use of multi-grade teaching (Glick and Sahn, 2006) and the condition of classrooms etc. However, the effect of these variables on enrolment or attendance is mixed and it is not unreasonable to argue that some of these measures might matter more for explaining school achievement than attendance.

3.4 Data, Variables and Methodology

3.4.1 Data and Variables

The dependent variable of interest in the empirical analysis is a dummy variable equal to 1 if a child of primary school age was attending primary school in 2005-2006. With this particular round of the DHS, the date at which the survey was carried out is important; the 2006 wave for Benin took place between the 3rd of August and 18th November 2006.

[53] Handa (2002) is a notable exception; His study considered a range of factors related to school enrolment in Mozambique on both the demand and supply side

[54] Although there may be a greater degree of measurement error here, as population ratios are usually based on estimates.

However, the Beninese government completed the elimination of primary school fees for all in 2006. Specifically, this announcement was made on October 14th 2006, which not only falls in the middle of the survey period, but also happens to be after the school year had already started.[55] As a result, there may be children surveyed in August or September that were not attending primary school, that might have done so in the absence of fees.[56] Unfortunately, it is not possible at this stage to pinpoint whether or not this policy was implemented immediately, or indeed was put in place for the following academic year. Nonetheless, in order to avoid any potential effects of this policy change, the dependent variable used is a dummy equal to 1 if a child was attending primary school during the previous (October 2005 - July 2006) school year.[57] Even then, there may thus be some confusion on the part of respondents: if they were surveyed in August (and perhaps even September), they may understand the 'current year' to mean that just past, as opposed to that which was about to begin in October. Controls for the month in which the survey took place are included in order to attempt to correct for any misunderstanding on the part of the respondent in later survey months.

The DHS data is stratified as follows:

Benin (**1**)
↓
Departments (**12**)
↓
Communes (**77**)
↓
"Clusters" (Primary Sampling Unit [PSU]) (**750**)
↓
Households (**17,511**)
↓
Individuals (**90,650**)
↙ ↘
Not School Age (**73,556**) School Age (**17,094**)

Of the 90,650 individuals, 17,094 were of school age. Of these, 57% were attending primary school. Control variables are included at the individual, household and commune level; table 3.2 presents summary statistics of these.

At the individual level, controls are included for a child's gender, whether (s)he was

[55] The school year runs from October to July in Benin
[56] Information regarding the cost of primary school fees is unfortunately unavailable at either the national or regional level.
[57] The DHS surveys ask not only "Did [the household member] attend school during the current year?" but also "Did [the household member] attend school during the previous year?" Responses to the latter question are used here.

adopted and whether or not (s)he worked alongside studying. Just under half of the sample was female (48%), 3.5% were adopted and it appears common that children of school age also work; some 73% did so.

Turning to the household level, a set of dummy variables is included in order to control for the household head's religion, the DHS wealth index (a composite index constructed using principal components analysis ranking households from 1 ['Poorest'] to 5 ['Richest']) and the education level of the household head (a set of dummy variables for none, primary, secondary or tertiary).[59]

Table 3.2 *Summary Statistics*

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
'Net Attendance 2005-06	0.570	0.495	0	1
Individual Level				
Gender (=1 if female)	0.475	0.499	0	1
Adopted	0.035	0.183	0	1
Worked	0.729	0.444	0	1
Worked in the field	0.318	0.444	0	1
Domestic Work	0.367	0.482	0	1
Other[58]	0.044	0.206	0	1
Household Level				
Household Religion				
Catholic	0.254	0.435	0	1
Protestant	0.057	0.232	0	1
Other Christian	0.070	0.255	0	1
Celeste	0.051	0.221	0	1
Islam	0.256	0.437	0	1
Vodoun	0.212	0.408	0	1
Other Traditional	0.033	0.179	0	1
Other Religion	0.013	0.115	0	1
No Religion	0.054	0.227	0	1
Household Wealth				
Poorest	0.222	0.416	0	1
Poorer	0.220	0.414	0	1
Middle	0.205	0.404	0	1
Richer	0.193	0.395	0	1
Richest	0.160	0.366	0	1
Household Head Education Level				
Primary	0.214	0.410	0	1
Secondary	0.130	0.337	0	1
Higher	0.018	0.132	0	1
School considered essential?	0.790	0.407	0	1
Household Size	8.174	4.104	2	36
Rural	0.639	.480	0	1
Commune Level				
Distance to School	1.279	0.971	0.134	5.399
(log) Schools per 5- 14 year olds	-5.977	0.238	-6.577	-5.471
Observations: 17,094				

[58] 'Other' includes categories such as *Construction* or *Auto Mechanic*.
[59] See DHS (2004) for a detailed report on the construction of the wealth index.

One quarter of households were Catholic, 26% Islamic and 21% Vodoun. Around 65% of household heads had no formal education whatsoever; of the remainder that did attend school, 21% had primary, 13% had secondary and the remaining 2% had a university education. Whilst ideally the child's mother or father's education level would be included in the analysis here (as the head is not always the parent of the children included in the sample), it was not possible to identify the parent of each child from the 2006 Beninese DHS due to missing data. *School Essential* is a dummy variable taking the value equal to one if the respondent answered yes to the survey question "Do you need to be able to send children to school".[60] This variable might reasonably be expected to capture household stated preferences for education – some 79% of households considered school to be essential. The average household size was just over 8 and this ranged from 2 to 36. Also included is a dummy variable equal to one if the child resided in a rural area (64% did so).

At the commune level, controls are included for the average distance to school and the (log) number of schools per school-age children in each commune.[61] To construct a measure of average distance to school, I follow Huisman and Smits (2009) as follows

$$Avg\ Distance = \frac{\sqrt{\frac{km^2}{No\ of\ Schools}}}{\pi} \qquad [4]$$

One unavoidable weakness of this approach lies with those children who reside close to the border of one commune, but might attend a school situated in another. However, this is perhaps the best approximation available, given the data on hand.[62]

3.4.2 Methodology

Whilst geographical variation in attendance rates is explored in detail below in the form of a multilevel model, a simple logistic regression is presented first. The advantage of doing this lies in the ability to compute average marginal effects which allow us to gain an understanding of the relative magnitudes of the covariates considered. The single level logistic regression estimated takes the following form

[60] The set of possible answers was 'No', 'Yes, essential' or 'Yes, more or less necessary'

[61] Unfortunately, it was not possible to obtain population estimates for the age range of 6-11, the official primary school age. Thus the number of schools per children aged 5 – 14 is shown.

[62] There are sometimes concerns surrounding endogeneity of distance-to-school variables: for example, the endogenous placement of schools, in areas where enrolment is low. However, this is unlikely to be a concern here: Firstly, the data is cross-sectional and therefore only measures the number of schools at one point in time – any phenomenon where more schools were being built as a result of initially low attendance would not be captured here. Secondly, the dependent variable is measured at the individual level and the distance variable at the commune level, so reverse causality is not likely to be an issue.

$$ln\left[\frac{p_i}{1-p_i}\right] = \beta_{0i} + \beta_1 \boldsymbol{S}_i + \beta_2 \boldsymbol{H}_i + \beta_2 \boldsymbol{C}_i + e_i \quad [5]$$

where the dependent variable is a dummy equal to 1 if child *i* was of official school age and attended school during the 2005-06 school year. **S** is the vector of student-level characteristics; **H** the vector of household level characteristics and the vector **C** contains commune-level characteristics.

3.5 Results

Table 3.3 presents benchmark results for equation [5]. In all following tables, the dependent variable is as outlined above. It can be considered a close approximation to the NER, although is again perhaps best defined as a net *attendance* rate. Average marginal effects are shown.[63] Column 1 includes the entire sample, whilst columns 2 and 3 perform the same regression for boys and girls respectively. Looking first at the individual level factors, the likelihood of being enrolled in primary school is significantly lower for girls (*Gender* =1 if the child was female) than it is for boys: this is in line with expectations, given the considerable gender disparities that exist(ed) in enrolment at primary school for Benin's young girls.

Specifically, the results suggest that girls aged 6-11 were around 9% less likely to attend primary school than are boys. Age and its square suggest a non-linear relationship with the dependent variable, reflecting the pattern outlined in figure 3.4 above. The results also suggest that adopted children are less likely to attend school than those related to the family; in a context where income might be low it is understandable that households might give preference to the education of biological children. Interestingly, however, column three highlights that the significance of this variable comes entirely from girls who are around 20% less likely to be enrolled in school than a biological daughter; adopted sons appear to face no such disadvantage compared to biological ones. Further inquiry showed that of four types of child (biological son, biological daughter, adopted son and adopted daughter), adopted daughters were the least likely of any to attend school, even when compared to adopted sons.[64] An alternative way to frame this inquiry is to include an interaction term (*gender * adopted*) in the full sample. When tested, this yielded an identical result, with a similar marginal effect (0.19). Lincove (2012) found a similar result in Uganda, however

[63] The same argument applies here as in chapter 2. The large no. of dummy variables means that it is most appropriate to present AMEs. See Bartus (2005) for further explanation.
[64] Results not shown, but available upon request

her results actually showed a larger impact for fostered boys than fostered girls – the opposite of what is presented here, whilst Huisman and Smits (2009) found an overall negative effect on the likelihood of foster children attending school in their panel of 30 developing countries. *Worked* is a dummy variable equal to 1 if children in the sample carried out any kind of work alongside studying. Again, girls appear to be at a disadvantage compared to boys, with the likelihood of being in school 5% lower for those girls who worked, compared to those who did not. This is explored in more detail below.

Table 3.3 *Results for full sample and by gender*

	All	*Male*	*Female*
	1	2	3
Individual level Characteristics			
Gender	-0.0879***		
	(0.0114)		
Age	0.481***	0.490***	0.469***
	(0.0240)	(0.0265)	(0.0360)
Age²	-0.0252***	-0.0253***	-0.0249***
	(0.00150)	(0.00157)	(0.00227)
Adopted	-0.134***	-0.0160	-0.203***
	(0.0368)	(0.0330)	(0.0456)
Worked	-0.0232	0.00301	-0.0477**
	(0.0157)	(0.0168)	(0.0186)
Household level Characteristics			
Household size	-0.00288**	-0.00328**	-0.00260
	(0.00123)	(0.00137)	(0.00172)
Religion			
Christian	-	-	-
Islam	-0.0728***	-0.0888***	-0.0567**
	(0.0196)	(0.0216)	(0.0228)
Traditional / Other	-0.0404***	-0.0457***	-0.0379**
	(0.0106)	(0.0130)	(0.0152)
Household Wealth level			
Poorest	-0.287***	-0.316***	-0.283***
	(0.0187)	(0.0232)	(0.0250)
Poorer	-0.197***	-0.247***	-0.171***
	(0.0165)	(0.0218)	(0.0201)
Middle	-0.107***	-0.169***	-0.0671***
	(0.0147)	(0.0216)	(0.0187)
Richer	-0.0314*	-0.0922***	0.00461
	(0.0164)	(0.0217)	(0.0191)
Richest	-	-	-
Household head's education			
None	-	-	-
Primary	0.0733***	0.0639***	0.0798***
	(0.0114)	(0.0156)	(0.0159)
Secondary	0.130***	0.125***	0.136***
	(0.0126)	(0.0168)	(0.0177)
Tertiary	0.0705*	0.170***	0.0431
	(0.0410)	(0.0356)	(0.0572)
School Considered essential?	0.00772	0.00398	0.0127
	(0.0149)	(0.0158)	(0.0202)
Rural	-0.0205	-0.0297**	-0.0139
	(0.0127)	(0.0144)	(0.0158)
Commune level Characteristics			
Distance to School	-0.0163	-0.0322***	0.00153
	(0.0104)	(0.0105)	(0.0131)
(ln) Schools per 5-14 year olds	0.159***	0.152***	0.165***
	(0.0443)	(0.0413)	(0.0561)
Observations	17,094	8,969	8,125

*Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01; Average Marginal Effects shown. Controls for the survey month included but not shown.*

Turning to the household's religion, the results suggest that children of Islamic households were around 7% less likely to be sent to school than those from Christian homes (the reference category).[65] Those of parents following *traditional / other* religion were about 4% less likely to attend. The DHS surveys also inquired as to the ethnicity of individuals, however this was often closely correlated to religion and did not provide any additional insights when tested in the model (not shown).

It is clear that as household wealth increases, so too does the likelihood that children attend primary school. For instance, those in the lowest wealth quintile were some 29% less likely to attend school than those in the richest (the reference category). This echoes results in studies such as Huisman and Smits (2009), or Delprato and Sabetes (2015) (Nigeria), which also found an increasing likelihood of school attendance as household wealth level increased, although Lincove (2012) found no effects of household wealth on school attendance in Uganda. The effects of household wealth again suggest some difference by gender: girls from the poorest households face a slightly lower disadvantage to boys, compared to the upper quintile, perhaps a reflection of the higher value placed on boys' work. This is explored in more detail below.

Similarly, household heads that had attended primary or secondary school were more likely to send their own children to primary school than those with no education (the reference category); there is little difference here by gender. Obviously it might be the case that richer families are often more educated, or it might be the case that Christian families are more likely to be wealthier (happen to reside in rich regions, etc.). In order to investigate, tables 3.4a and 3.4b display the predictive margins of being in a Christian home by wealth level and the predictive margins of having a more educated household head by wealth level respectively.

It is clear that, holding wealth level constant, household religion and household head's education still have an impact on the likelihood of a child attending school. In particular, children of Christian parents are consistently around 5-6 % more likely to attend school, regardless of household wealth level. Turning to table 3.4b, again it is clear that holding wealth constant, the likelihood that a child is sent to school is greater for those household heads with primary compared to no education and greater for those with secondary compared to primary education. Interestingly, holding wealth level fixed, the household

[65] The 'Christian' category includes 'Protestant', 'Catholic', 'Celsete' & 'Other Christian.

head having a tertiary education does not increase the probability that a child will attend school compared to a secondary level education.

Table 3.4a *Predictive margins of religion, by household wealth level.*

Wealth Level	Christian	Predictive Margin
Poorest	No	0.394
Poorest	Yes	0.455
Poorer	No	0.489
Poorer	Yes	0.551
Middle	No	0.585
Middle	Yes	0.644
Richer	No	0.655
Richer	Yes	0.710
Richest	No	0.675
Richest	Yes	0.728

Table 3.4b *Predictive margins of household head's education, by household wealth level.*

Wealth Level	Head Education	Predictive Margin
Poorest	None	0.383
Poorest	Primary	0.464
Poorest	Secondary	0.528
Poorest	Tertiary	0.446
Poorer	None	0.478
Poorer	Primary	0.561
Poorer	Secondary	0.623
Poorer	Tertiary	0.543
Middle	None	0.575
Middle	Primary	0.655
Middle	Secondary	0.710
Middle	Tertiary	0.637
Richer	None	0.647
Richer	Primary	0.720
Richer	Secondary	0.769
Richer	Tertiary	0.704
Richest	None	0.667
Richest	Primary	0.738
Richest	Secondary	0.785
Richest	Tertiary	0.723

Looking at the commune-averaged variables in table 3.3, male children in rural areas are less likely to be enrolled than males in urban areas; it also seems that distance to school matters only for boys. This result emerges in direct conflict to the theory that parents will be less likely to send their female children to school because of the long distance to walk, fear of attack etc. The supply of schools is however important – the variable *Schools per 5-14 year olds* reflects the number of schools in a given commune, divided by the population

aged 5 to 14 years.[66] The positive and significant coefficient on this variable for children of both genders suggests that having controlled for demand-side factors, the number of schools also matters for attendance rates. However, these results warrant closer inspection; table 3.5 looks further at the role of school supply. Specifically, due to the relatively high correlation between average distance to school and number of schools (-0.57), the effects of these variables are examined in isolation.

The results in table 3.5 suggest that for the complete, rural or urban sample, both distance and the number of schools appear to affect the likelihood that a child is enrolled. However, the magnitude and level of statistical significance depends on the exact specification. In an urban setting, distance is barely significant. This is not surprising; the average distance to school in urban areas would be shorter than in rural ones, so the usual arguments about the opportunity cost of the time travelling do not apply to the same extent as in rural areas. However, the number of schools per children aged 5-14 is a significant explanatory factor. In a rural setting, both distance and the number of schools are significant determinants of whether a child will attend school.

Returning to the result in column 3, table 3.3, that suggested distance was not a significant explanatory factor for girls' school attendance, the effects of both gender and location are investigated in table 3.6. Immediately clear is that the number of schools in a commune is a significant explanatory factor for all groups, although only at the 10% level for girls in urban areas. Again, it is confirmed that the distance to school is not a significant explanatory variable for girls in either rural or urban areas. However, it does seem to matter for boys. Distance has frequently been used as a proxy for school costs in enrolment / attendance studies: for example, Lincove (2012) [Uganda] and Handa (2002) [Mozambique] found negative effects of distance to school and travel time to school respectively on school attendance.

[66] This age range was chosen as *INSAE* produce population estimates for ages 5-9 and 10-14. It was deemed preferable to use these figures as published, rather than construct an age range of 6-11, which would require use of arbitrary estimates for population growth.

Table 3.5 *School supply characteristics; Rural v Urban.*

	Full Sample			***Urban***			***Rural***		
	1	2	3	4	5	6	7	8	9
Individual level Characteristics									
Gender	-0.0879***	-0.0882***	-0.0877***	-0.0994***	-0.100***	-0.0992***	-0.0797***	-0.0797***	-0.0795***
	(0.0114)	(0.0116)	(0.0114)	(0.0197)	(0.0201)	(0.0196)	(0.0105)	(0.0105)	(0.0106)
Age	0.481***	0.484***	0.481***	0.514***	0.515***	0.513***	0.472***	0.476***	0.472***
	(0.0240)	(0.0242)	(0.0237)	(0.0371)	(0.0384)	(0.0369)	(0.0289)	(0.0288)	(0.0290)
Age^2	-0.0252***	-0.0254***	-0.0252***	-0.0282***	-0.0283***	-0.0282***	-0.0241***	-0.0243***	-0.0241***
	(0.00150)	(0.00152)	(0.00148)	(0.00245)	(0.00254)	(0.00244)	(0.00168)	(0.00167)	(0.00168)
Adopted	-0.134***	-0.134***	-0.136***	-0.181***	-0.179***	-0.182***	-0.0920***	-0.0924***	-0.0934***
	(0.0368)	(0.0373)	(0.0361)	(0.0496)	(0.0512)	(0.0476)	(0.0277)	(0.0279)	(0.0276)
Worked	-0.0232	-0.0242	-0.0247	-0.0438**	-0.0420**	-0.0446**	-0.00712	-0.0105	-0.00915
	(0.0157)	(0.0158)	(0.0155)	(0.0208)	(0.0206)	(0.0203)	(0.0173)	(0.0176)	(0.0172)
Household level Characteristics									
Household size	-0.00288**	-0.00336***	-0.00321**	-0.00471**	-0.00471**	-0.00491**	-0.00194	-0.00273*	-0.00231
	(0.00123)	(0.00126)	(0.00126)	(0.00197)	(0.00190)	(0.00196)	(0.00155)	(0.00164)	(0.00154)
Religion									
Islam	-0.0728***	-0.0776***	-0.0842***	-0.0522**	-0.0547**	-0.0562**	-0.0830***	-0.0900***	-0.102***
	(0.0196)	(0.0213)	(0.0186)	(0.0259)	(0.0263)	(0.0252)	(0.0246)	(0.0284)	(0.0222)
Traditional / Other	-0.0404***	-0.0380***	-0.0371***	-0.000785	0.000264	0.000934	-0.0573***	-0.0540***	-0.0526***
	(0.0106)	(0.0115)	(0.0111)	(0.0192)	(0.0190)	(0.0195)	(0.0127)	(0.0146)	(0.0130)
Household Wealth Level									
Poorest	-0.287***	-0.288***	-0.298***	-0.289***	-0.288***	-0.297***	-0.315***	-0.322***	-0.321***
	(0.0187)	(0.0190)	(0.0183)	(0.0267)	(0.0263)	(0.0254)	(0.0262)	(0.0276)	(0.0260)
Poorer	-0.197***	-0.202***	-0.206***	-0.192***	-0.194***	-0.198***	-0.228***	-0.238***	-0.230***
	(0.0165)	(0.0175)	(0.0162)	(0.0254)	(0.0251)	(0.0255)	(0.0251)	(0.0267)	(0.0247)
Middle	-0.107***	-0.111***	-0.115***	-0.109***	-0.110***	-0.114***	-0.135***	-0.145***	-0.136***
	(0.0147)	(0.0161)	(0.0145)	(0.0221)	(0.0225)	(0.0210)	(0.0228)	(0.0248)	(0.0226)
Richer	-0.0314*	-0.0340*	-0.0362**	-0.0179	-0.0173	-0.0207	-0.0749***	-0.0842***	-0.0744**
	(0.0164)	(0.0179)	(0.0157)	(0.0180)	(0.0178)	(0.0178)	(0.0290)	(0.0307)	(0.0290)
Household head's education									
Primary	0.0733***	0.0786***	0.0742***	0.0818***	0.0831***	0.0823***	0.0645***	0.0726***	0.0656***
	(0.0114)	(0.0109)	(0.0114)	(0.0161)	(0.0157)	(0.0163)	(0.0143)	(0.0141)	(0.0142)
Secondary	0.130***	0.134***	0.131***	0.107***	0.108***	0.108***	0.158***	0.165***	0.159***
	(0.0126)	(0.0130)	(0.0127)	(0.0144)	(0.0143)	(0.0144)	(0.0212)	(0.0214)	(0.0213)
Tertiary	0.0705*	0.0722*	0.0729*	0.0581	0.0570	0.0591	0.277***	0.292***	0.276***
	(0.0410)	(0.0400)	(0.0410)	(0.0429)	(0.0426)	(0.0421)	(0.0800)	(0.0752)	(0.0805)
School Essential?	0.00772	0.0168	0.00976	0.0126	0.0214	0.0119	0.00262	0.0112	0.00787
	(0.0149)	(0.0139)	(0.0150)	(0.0167)	(0.0170)	(0.0169)	(0.0208)	(0.0188)	(0.0211)
Rural	-0.0205	-0.0235*	-0.0205	-	-	-	-	-	-
	(0.0127)	(0.0137)	(0.0126)						
Commune - level Characteristics									
Distance to School	-0.0163	-0.0341***		-0.00881	-0.0234*		-0.0230*	-0.0413***	
	(0.0104)	(0.0102)		(0.0152)	(0.0129)		(0.0134)	(0.0140)	
(ln) Schools per 5-14 year olds	0.159***		0.182***	0.108*		0.126***	0.183***		0.209***
	(0.0443)		(0.0394)	(0.0561)		(0.0448)	(0.0515)		(0.0479)
Observations	17,094	17,094	17,094	6,171	6,171	6,171	10,923	10,923	10,923

*Standard errors, clustered at the commune level, in parentheses; *p<0.1; **p<0.05; ***p<0.01;* Controls for the survey month included but not shown

Table 3.6 *School Supply Characteristics; Rural vs Urban by gender.*

	Rural				***Urban***			
	Male	Female	Male	Female	Male	Female	Male	Female
	1	2	3	4	5	6	7	8
Individual level Characteristics								
Age	0.498***	0.441***	0.497***	0.436***	0.489***	0.531***	0.479***	0.532***
	(0.0358)	(0.0413)	(0.0352)	(0.0415)	(0.0462)	(0.0618)	(0.0448)	(0.0607)
Age²	-0.0253***	-0.0226***	-0.0253***	-0.0223***	-0.0262***	-0.0298***	-0.0256***	-0.0298***
	(0.00209)	(0.00239)	(0.00206)	(0.00241)	(0.00285)	(0.00400)	(0.00275)	(0.00394)
Adopted	-0.0260	-0.145***	-0.0282	-0.144***	0.000112	-0.258***	-0.00952	-0.257***
	(0.0406)	(0.0356)	(0.0407)	(0.0351)	(0.0462)	(0.0602)	(0.0446)	(0.0587)
Worked	0.0105	-0.0337*	0.00957	-0.0311	-0.0139	-0.0621**	-0.0156	-0.0642***
	(0.0198)	(0.0198)	(0.0193)	(0.0198)	(0.0251)	(0.0256)	(0.0243)	(0.0248)
Household level Characteristics								
Household Size	-0.00255	-0.00280	-0.00280*	-0.00184	-0.00529**	-0.00445*	-0.00612**	-0.00400*
	(0.00166)	(0.00228)	(0.00161)	(0.00213)	(0.00235)	(0.00238)	(0.00248)	(0.00239)
Religion								
Islam	-0.110***	-0.0658**	-0.131***	-0.0671**	-0.0687**	-0.0489*	-0.0777**	-0.0391
	(0.0289)	(0.0334)	(0.0222)	(0.0288)	(0.0312)	(0.0263)	(0.0307)	(0.0245)
Traditional / Other	-0.0539***	-0.0541***	-0.0509***	-0.0549***	-0.0204	0.0154	-0.0141	0.0105
	(0.0170)	(0.0198)	(0.0163)	(0.0174)	(0.0225)	(0.0252)	(0.0244)	(0.0244)
Household Wealth Level								
Poorest	-0.342***	-0.306***	-0.345***	-0.303***	-0.271***	-0.331***	-0.299***	-0.319***
	(0.0333)	(0.0397)	(0.0341)	(0.0379)	(0.0332)	(0.0354)	(0.0320)	(0.0338)
Poorer	-0.274***	-0.206***	-0.267***	-0.197***	-0.208***	-0.202***	-0.228***	-0.186***
	(0.0332)	(0.0380)	(0.0338)	(0.0358)	(0.0321)	(0.0305)	(0.0324)	(0.0310)
Middle	-0.187***	-0.108***	-0.181***	-0.0971***	-0.147***	-0.0844***	-0.167***	-0.0724***
	(0.0324)	(0.0375)	(0.0331)	(0.0346)	(0.0311)	(0.0297)	(0.0297)	(0.0276)
Richer	-0.104***	-0.0666*	-0.0970**	-0.0556	-0.0790***	0.0211	-0.0903***	0.0273
	(0.0391)	(0.0404)	(0.0393)	(0.0373)	(0.0258)	(0.0211)	(0.0252)	(0.0197)
Household head's education								
Primary	0.0713***	0.0740***	0.0658***	0.0655***	0.0636***	0.0980***	0.0642**	0.0954***
	(0.0169)	(0.0200)	(0.0165)	(0.0205)	(0.0244)	(0.0169)	(0.0249)	(0.0172)
Secondary	0.148***	0.183***	0.142***	0.177***	0.105***	0.109***	0.104***	0.108***
	(0.0274)	(0.0298)	(0.0266)	(0.0305)	(0.0216)	(0.0168)	(0.0215)	(0.0170)
Tertiary	0.273*	0.300***	0.249	0.286***	0.136***	0.0374	0.140***	0.0375
	(0.159)	(0.104)	(0.171)	(0.108)	(0.0339)	(0.0607)	(0.0328)	(0.0599)
School Considered essential?	0.0123	0.0124	0.0110	0.00624	0.0106	0.0284	0.00186	0.0187
	(0.0204)	(0.0248)	(0.0231)	(0.0263)	(0.0194)	(0.0220)	(0.0201)	(0.0213)
Commune level Characteristics								
Distance to School	-0.0539***	-0.0262			-0.0430***	-0.00113		
	(0.0147)	(0.0162)			(0.0133)	(0.0159)		
(ln) Schools per 5-14 year olds			0.226***	0.187***			0.151***	0.102*
			(0.0514)	(0.0565)			(0.0489)	(0.0564)
Observations	5,792	5,131	5,792	5,131	3,177	2,994	3,177	2,994

*Standard errors, clustered at the commune level, in parentheses; *p<0.1; **p<0.05; ***p<0.01;* Controls for the survey month included but not shown

As a further check on the role of distance to school, table 3.7 investigates the extent to which this variable can shed light on the opportunity cost of attendance by subdividing the sample by gender, whether or not children worked and the type of work carried out.[67] Again, there is no significant effect of distance on girls' attendance if they worked (column 6), however the result for boys presents strong evidence that the distance to school is a real concern for those that worked alongside studying (column 2). This supports the hypothesis that the opportunity cost of schooling has been taken into account when deciding whether or not to send a child to school. For those that did not work (column 1), distance is not a significant explanatory factor. A number of explanations for the differences observed between genders might be considered:

Firstly, boys' labour, or the nature of their labour, may contribute more to family income than girls' and so the opportunity costs of travel time are more heavily felt. Secondly, the climate in Benin at that time of promoting enrolment for all might have seen parents under pressure to (be seen to) send daughters to school, so even those required to work were encouraged to attend to a greater extent than were sons.

Columns 3, 4, 7 and 8 of table 3.7 examine the effect of distance, with the sample divided by the type of work carried out by boys and girls respectively.[68] It is clear that those engaging in *Work in the Field* faced the greatest opportunity cost to attending school; an increase in average distance to school of 1km reduced the likelihood that boys would attend school by around 8% on average. Interestingly, it appears that girls who worked in the field also faced a higher opportunity cost to attending school, but this is around half of the magnitude of boys (~4.5%) and only statistically significant at the 5% level. Boys that undertook domestic work were around 4% less likely to attend school as the average distance increased by 1 km. No such effect is seen for girls.

[67] No significant differences were uncovered here when differentiating between rural and urban, or when including the 'Number of Schools' variable.

[68] These are the two main types of work carried out by school-age children; Other types included 'Construction', 'Auto Mechanic' or ' subterraneous work', however these categories only accounted for a small number of children; there were insufficient observations to run a similar logistic model.

Table 3.7 *Marginal effects of (commune) average distance to school, by gender and work status*

	Boys				Girls			
	Didn't work	***Worked***	***In field***	***Domestic***	***Didn't work***	***Worked***	***In field***	***Domestic***
	1	2	3	4	5	6	7	8
Individual level Characteristics								
Age	0.525***	0.445***	0.430***	0.370***	0.551***	0.424***	0.353***	0.450***
	(0.0453)	(0.0324)	(0.0572)	(0.0565)	(0.0738)	(0.0394)	(0.0703)	(0.0546)
Age²	-0.0265***	-0.0231***	-0.0215***	-0.0186***	-0.0295***	-0.0224***	-0.0176***	-0.0236***
	(0.00286)	(0.00188)	(0.00319)	(0.00359)	(0.00481)	(0.00233)	(0.00395)	(0.00335)
Adopted	0.0581	-0.0554	-0.0712	-0.0806	-0.214***	-0.194***	-0.127**	-0.214***
	(0.0538)	(0.0370)	(0.0503)	(0.0496)	(0.0588)	(0.0464)	(0.0638)	(0.0465)
Household level Characteristics								
Household Size	-0.00627***	-0.00229	-0.00179	-0.00149	-0.00846**	-0.00186	-0.00342	0.000311
	(0.00232)	(0.00162)	(0.00196)	(0.00234)	(0.00346)	(0.00231)	(0.00306)	(0.00240)
Household Religion								
Islam	-0.0395	-0.119***	-0.120***	-0.0761**	-0.0127	-0.0786***	-0.170***	-0.0753***
	(0.0333)	(0.0251)	(0.0284)	(0.0309)	(0.0332)	(0.0290)	(0.0550)	(0.0261)
Traditional / Other	-0.0534**	-0.0451**	-0.0642***	-0.0462*	-0.0691***	-0.0290	-0.109***	0.0138
	(0.0241)	(0.0183)	(0.0204)	(0.0239)	(0.0235)	(0.0191)	(0.0251)	(0.0240)
Household Wealth Level								
Poorest	-0.279***	-0.327***	-0.393***	-0.245***	-0.263***	-0.280***	-0.551***	-0.279***
	(0.0453)	(0.0354)	(0.0660)	(0.0434)	(0.0416)	(0.0262)	(0.103)	(0.0299)
Poorer	-0.233***	-0.259***	-0.330***	-0.179***	-0.208***	-0.159***	-0.433***	-0.163***
	(0.0419)	(0.0350)	(0.0648)	(0.0378)	(0.0385)	(0.0248)	(0.105)	(0.0225)
Middle	-0.163***	-0.173***	-0.217***	-0.130***	-0.123***	-0.0531**	-0.342***	-0.0381
	(0.0376)	(0.0351)	(0.0603)	(0.0381)	(0.0369)	(0.0229)	(0.102)	(0.0247)
Richer	-0.0700*	-0.110***	-0.183***	-0.0703*	-0.0393	0.0171	-0.308***	0.0456*
	(0.0366)	(0.0354)	(0.0619)	(0.0386)	(0.0289)	(0.0247)	(0.105)	(0.0242)
Household head education level								
Primary	0.0576**	0.0728***	0.108***	0.0116	0.0735***	0.0908***	0.0697**	0.0937***
	(0.0261)	(0.0160)	(0.0234)	(0.0211)	(0.0273)	(0.0176)	(0.0310)	(0.0205)
Secondary	0.121***	0.137***	0.188***	0.0487	0.134***	0.142***	0.111*	0.127***
	(0.0271)	(0.0277)	(0.0390)	(0.0322)	(0.0280)	(0.0229)	(0.0668)	(0.0262)
Tertiary	0.219***	0.124**	-	-0.00381	0.277***	-0.0714	-	-0.0682
	(0.0497)	(0.0567)		(0.0528)	(0.0540)	(0.0468)		(0.0552)
School Essential?	0.00477	0.0103	0.0161	0.00639	0.0226	0.0208	0.0135	-0.00475
	(0.0239)	(0.0206)	(0.0248)	(0.0247)	(0.0273)	(0.0225)	(0.0421)	(0.0203)
Rural	-0.0458*	-0.0250	-0.0188	-0.0307	-0.0424**	-0.00995	0.0169	-0.0140
	(0.0235)	(0.0178)	(0.0201)	(0.0221)	(0.0213)	(0.0204)	(0.0362)	(0.0212)
Commune level Characteristics								
Distance to School	-0.0123	-0.0615***	-0.0786***	-0.0442***	-0.00863	-0.0185	-0.0444**	-0.0193
	(0.0154)	(0.0119)	(0.0156)	(0.0132)	(0.0161)	(0.0135)	(0.0202)	(0.0149)
Observations	2,562	6,407	3,702	2,328	2,063	6,062	1,738	3,941

*Standard errors, clustered at the commune level, in parentheses; *p<0.1; **p<0.05; ***p<0.01; Controls for the survey month included but not shown. 'Other work' not shown due to small sample size.*

The fact that the average marginal effect of average distance to school for boys who *work in the field* is almost double that of girls (and more strongly significant) suggests that the opportunity cost of attending school is higher for them. We see no significant effects of distance for the subsample of boys and girls that did not work. This result echoes the findings of, for example, Colclough *et al.* (2000) who found that boys in both Guinea and Ethiopia that had dropped out of school did so primarily to earn money. However, Buchmann (2000) found that from a sample of 146 children who had dropped out of school in Kenya, only one did so for employment and one to help in the household.

3.5.1 Testing an alternative dependent variable

Whilst the dependent variable used above adheres closely to a measure of net enrolment in Benin, it is not a perfect barometer of primary school attendance. Figure 3.4 showed that many children in primary school in 2005-06 fell outside the official age category of 6-11. Furthermore, related studies in this field have imposed different criteria for classifying school attendance status: Huisman and Smits (2009) consider only those aged 8-11 in a panel of 30 countries; Lincove (2015), those aged 6-13 in Nigeria and Uganda (although crucially, omit any children that had already progressed to secondary school); Buchmann (2000) those aged 13-18 in Kenya; Lincove (2012), those aged 6-12 years old in Uganda. To the best of my knowledge, only Deininger (2003) provides estimates for a number of alternative age ranges [Uganda].

Whilst some of this disparity in the dependent variable used naturally results from different official starting ages and length of school cycles in each country, it can hamper the comparability of results across studies and across countries. An appealing approach is outlined in fhi360 (2013), who argue in favour of classing the school age as between 7 and 14 years old and abandoning the arbitrary definitions of 'primary' or 'secondary'. These bounds are influenced by the fact that "…in all countries, compulsory education begins by age 7 or earlier" and "…the ILO minimum age convention establishes age 15 as the minimum legal age for entering any form of employment…" (fhi360, 2013:48). As a result, any child falling within this age bracket should be expected to be in school. To follow this definition for the Beninese sample would lead to the loss of those children aged 5 or 6 who are, by all accounts, of school age. However, by extending the upper age range of the sample to 14, the analysis can account for many more children attending, or who have completed, primary school. In the robustness checks included in Appendix C, the dependent variable is thus a dummy equal to 1 if (i) a child is aged between 5 and 14 and (ii) (s)he is

either in or has completed primary school.[69] All of the previously reported results hold using this alternative dependent variable; the only differences arising are small changes in the magnitude of certain independent variables. For the sake of brevity, only a replication of table 3.3 is included.

3.6 Toward a multilevel approach

The results in section 3.5 have highlighted that factors on both the demand and supply side were important determinants of school attendance in Benin in the 2005-06 school year. As shown in section 3.2, stark regional disparities exist Benin with regards to attendance rates. Whilst a number of studies acknowledge that factors at the household, community or district level might impact on school attendance in SSA (for example Lavy (1996), Handa (2005) or Huisman and Smits (2009)), only Delprato and Sabetes (2015) have explicitly modelled this econometrically by taking account of unobserved heterogeneity between higher level clusters. The consequence for the econometric analysis of ignoring this is a violation of the assumption that observations are independent from one another; unobserved heterogeneity at higher levels leads to cluster-level interdependence between units (Skrondal & Rabe-Hesketh, 2004). The traditional approach to dealing with such data is to turn to a multilevel (or hierarchical) linear model (MLM) (HLM), the simplest of which is the random-intercepts or *variance components* model which estimates a random intercept for every higher level unit, such as commune or department etc.

3.6.1 Random Intercepts Model

After testing various multilevel structures, it turned out that whilst all performed better than the single level model above, a three-level model, as shown below, was preferred to a any two-level model.[70]

The basic variance components model takes the form

$$ln\left[\frac{p_{ijkl}}{1-p_{ijkl}}\right] = \beta_0 + \mu_k + \varphi_{jk} + \varepsilon_{ijk} \qquad [6]$$

[69] Thus, children currently in secondary school and those who have completed primary school but not gone on to secondary education are also included.
[70] This judgement was made on the basis of LR-test statistics.

where

$\mu_k \sim N(0, \sigma_\mu^2)$

$\varphi_{jk} \sim N(0, \sigma_\varphi^2)$

$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$ [7]

$\ln\left[\frac{p_{ijk}}{1-p_{ijk}}\right]$ is the log-odds that child *i* in household *j* in commune *k* is attending school. β_0 is the intercept shared by all individuals, households and communes. μ_k is the effect of commune *k*, φ_{jk} the effect of household *j* and ε_{ijk} is the child level residual error.

The models here are computed using second-order penalized quasi-likelihood (PQL2) in *MLwiN* (Rabash *et al.*, 2015) via the Stata module *runmlwin* (Leckie and Charlton, 2013). When selecting the appropriate means by which to estimate multilevel equations, it is necessary to choose a method that is both the most unbiased, but also computationally feasible. Simulations in Rodriguez and Goldman (2001) show that 2nd order Penalised Quasi Likelihood (PQL) estimation provided the closest approximation to maximum likelihood estimation (MLE), out of the choice of 1st and 2nd order marginal quasi-likelihood (MQL) and 1st and 2nd order PQL. Whilst, ideally, MLE would be used to obtain all the estimates, this is computationally very intensive for models beyond the null; Stata's commands such as *xtmelogit* often take many hours or days to converge, if at all. Appendix D displays estimates of the variance of the random effects in equation [6] using MQL1, MQL2, PQL1, PQL2 and MLE; Whilst there is still a downward bias in the estimates of σ_μ^2 and σ_φ^2 compared to MLE, PQL2 performs substantially better than the other quasi-likelihood estimators. Given that MLE estimates for models beyond the null are computationally very difficult to obtain, PQL2 is the preferred method here.

Column 1 of table 3.8 displays results from the random intercepts model, as per equation [6]. Odds ratios are displayed, as it is not possible to compute marginal effects for multilevel models. Thus, it is difficult to compare the magnitude of the covariates with the single level logistic regressions. The estimate of β_0 suggests that the log-odds of a child of school age attending primary school in an 'average' household / commune is 0.32.

Table 3.8 *Random Intercepts and random slopes model*

		Random Intercepts			***Random Slopes***
	1	2	3	4	5
Individual level Characteristics					
Gender		-0.538***	-0.582***	-0.582***	-0.578***
		(0.042)	(0.043)	(0.043)	(0.043)
Age		3.171***	3.183***	3.182***	3.163***
		(0.139)	(0.140)	(0.140)	(0.139)
Age²		-0.167***	-0.167***	-0.167***	-0.166
		(0.008)	(0.008)	(0.008)	(0.008)
Adopted		-0.658***	-0.924***	-0.917***	-0.929***
		(0.117)	(0.118)	(0.118)	(0.117)
Worked		-0.297***	-0.167***	-0.156***	-0.137**
		(0.055)	(0.055)	(0.054)	(0.054)
Household level Characteristics					
Household size			-0.019***	-0.015**	-0.015***
			(0.007)	(0.007)	(0.007)
Religion					
Christian		-	-	-	-
Islam			-0.386***	-0.311***	-0.327***
			(0.078)	(0.079)	(0.079)
Traditional / Other			-0.401***	-0.414***	-0.413***
			(0.063)	(0.063)	(0.063)
Grand mean - centred wealth level			0.507***	0.497***	0.498***
			(0.024)	(0.024)	(0.033)
Household head's education					
None			-	-	-
Primary			0.460***	0.447***	0.329***
			(0.063)	(0.063)	(0.063)
Secondary			0.714***	0.701***	0.696***
			(0.089)	(0.089)	(0.089)
Tertiary			0.418**	0.368*	0.371*
			(0.207)	(0.206)	(0.206)
School Essential?			0.154**	0.134**	0.125***
			(0.064)	(0.064)	(0.064)
Rural			-0.248***		-0.236***
			(0.062)		(0.062)
Commune - level Characteristics					
Distance to School				-0.165**	-0.0210
				(0.079)	(0.077)
(ln) Schools per 5-14 year olds				1.073***	1.128
				(0.303)	(0.289)
Intercept (β_0)	0.322***	-13.528***	-13.544***	-6.646***	-6.422***
	(0.089)	(0.577)	(0.603)	(1.842)	(1.764)
Random Effects					
Level 2: Household					
Intercept variance	1.223	1.800	1.418	1.422	1.373
VPC_{φ}:	*0.240*	*0.310*	*0.279*	*0.287*	-
ICC:	*0.353*	*0.430*	*0.352*	*0.336*	-
Level 3: Commune					
Intercept variance	0.573	0.700	0.368	0.241	0.214
VPC_{μ}:	*0.113*	*0.120*	*0.073*	*0.049*	-
ICC:	*0.113*	*0.120*	*0.073*	*0.049*	-
Wealth Slope Variance	-	-	-	-	0.037
Intercept – Wealth Slope Covariance	-	-	-	-	-0.043
Observations:	17,094	17,094	17,094	17,094	17,094

*Odds ratios shown. Standard errors in parentheses; *p<0.1; **p<0.05;***p<0.01*

In order to examine and interpret each of the variance components outlined above in [7], a variance partition coefficient (VPC) can be calculated which, for the unconditional model in column (1) of table 3.8, "…report[s] the proportion of the observed variance at each level of the model hierarchy" (Leckie, 2013:21) Thus the VPC provides an indication of those levels at which most residual variation in the likelihood of attending school exists. The discussion in section 3.2 has already highlighted the significant inter-commune disparity in

school attendance rates, so the expectation is that a significant amount of variation will exist at this level. The VPC's are calculated for the commune and household respectively, as follows:

$$VPC_{\mu} = \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2+\sigma_{\varphi}^2+\sigma_{\varepsilon}^2} \quad [8]$$

$$VPC_{\varphi} = \frac{\sigma_{\varphi}^2}{\sigma_{\mu}^2+\sigma_{\varphi}^2+\sigma_{\varepsilon}^2} \quad [9]$$

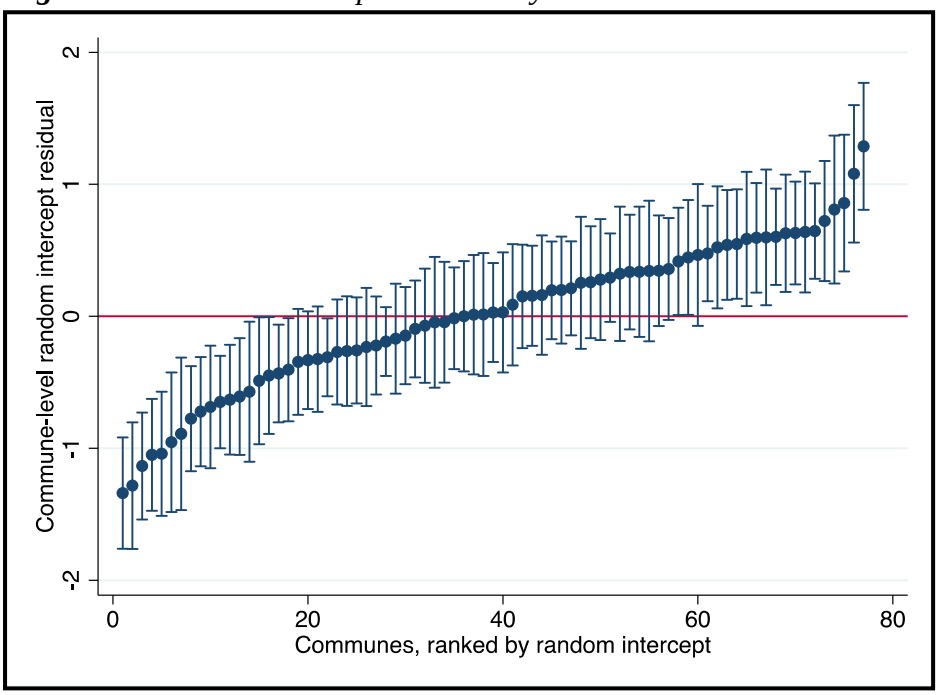
In the case of a binary outcome, σ_{ε}^2 is fixed at $\frac{\pi^2}{3} \approx 3.29$, the variance of the standard logistic distribution. For the null model, the VPC's are 0.240 and 0.113 for the household and commune respectively. Thus for the null model considered in column 1, 24% of variation in school attendance rates is between households, but only 11.3% between communes. An alternative means by which to interpret variance components is the intra-class correlation coefficient (ICC), which measures correlation or similarity between observed responses within a given higher-level cluster unit. The ICC for the commune level, ρ_{μ}, and the household level, ρ_{φ} is calculated as follows

$$\rho_{\mu} = \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2+\sigma_{\varphi}^2+\sigma_{\varepsilon}^2} \quad [10]$$

$$\rho_{\varphi} = \frac{\sigma_{\mu}^2+\sigma_{\varphi}^2}{\sigma_{\mu}^2+\sigma_{\varphi}^2+\sigma_{\varepsilon}^2} \quad [11]$$

Thus an ICC of 0.113 for the commune level represents the between-commune correlation in the odds that a child is attending school. An ICC of 0.353 shows that the between-household within-commune correlation is much higher, suggesting a higher correlation in the odds that any two children from different households in the one commune are attending school, compared to any two children from different households in different communes.

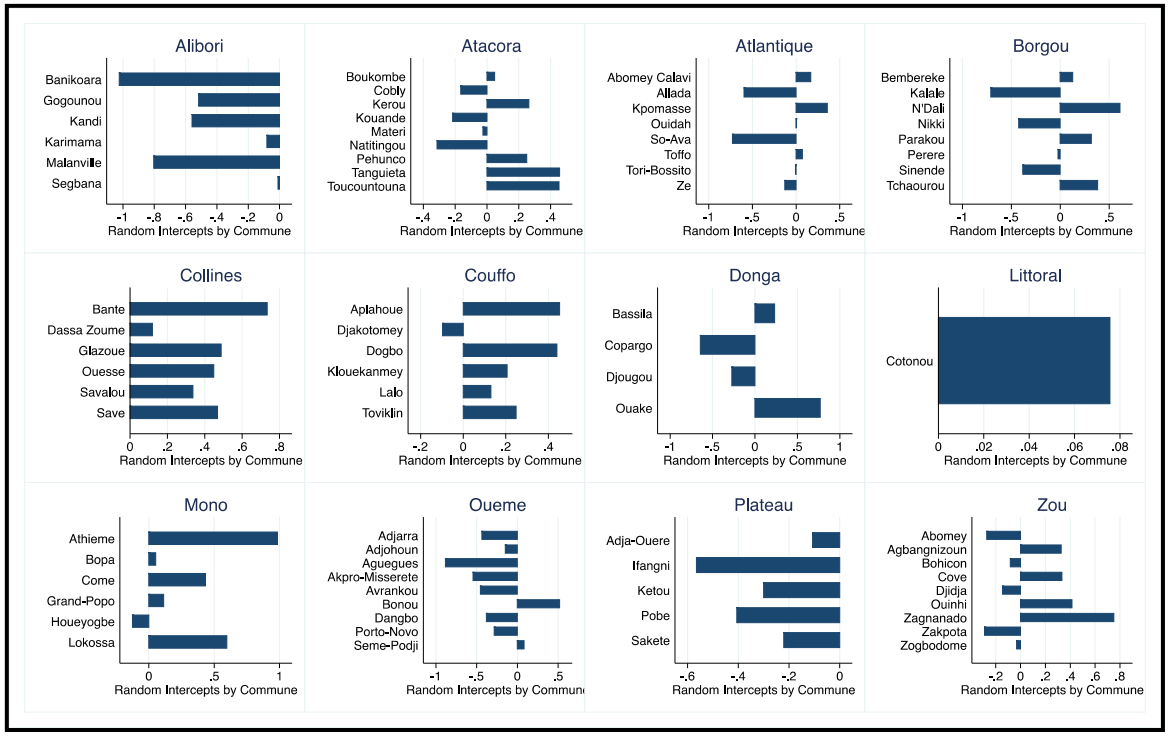
Figure 3.6 *Random intercept residuals by commune.*



A caterpillar plot of the commune-level residuals, in figure 3.6, shows a significant number of communes where the 95% confidence interval does not overlap with zero, suggesting that attendance rates are significantly higher or lower than average in these communes. The significant *between-commune-within-department* variance, discussed in section two, is also confirmed here; figure 3.7 illustrates that whilst some departments, such as Alibori or Plateau, contain only communes with a negative random intercept residual, others such as Atacora, Borgou or Zou, contain communes where the random effect is both above and below the average.[71]

[71] NB. Initial diagnostics deemed that the computational complexity of adding a department level to the model in Table 9 led to little gain in model performance. As such it was not included, but the graphical insight here is nonetheless useful.

Figure 3.7 *Between-commune within-department variation in random intercepts.*



In column 2 of table 3.8, covariates are included at the individual level. The VPC in columns 2 - 4, takes on a slightly different interpretation; in conditional models, it represents the degree of *unexplained* variance that exists at each higher level. Having controlled for individual level factors, we see that the 12% of residual variance in school attendance exists between communes and some 31% between households.

When household-level variables are included in column 3 of table 3.8, the variance component attributable to the household level falls from 1.800 to 1.418, representing a reduction in between-household variance of 21%.[72] Thus, the household covariates included (household wealth, religion, size and stated preferences for education) explain around a fifth of the the residual variation in attendance rates. The remaining household-level variation is attributable to some unobserved factors not accounted for here. The VPC for the household and commune fall to 27.9% and 7.3% respectively.

In column 4, commune-level variables are added to the model. Immediately clear is that the commune-level variance component falls from 0.355 to 0.241; the VPC for communes falls from 7% to 4.9%. Thus the inclusion of distance to school and number of schools in the model helps to explain around one third of the commune-level variation in primary

[72] In the multilevel models here, mean-centred wealth is included rather than the individual wealth quintiles, allowing estimation of the random slopes model, below.

school attendance. This result suggests that the regional differences observed in Benin are down to other unobserved factors at the commune-level. These might be, for example, regional differences in labour markets, culture or traditions. The model has also shown that overall, relatively little of the regional variation in attendance rates displayed in Benin is attributable to commune-level factors: Only 4.9% of the total remaining variation is at the commune-level, whilst some 30% is due to factors at the household level. Thus the greatest improvements in attendance rates might be realized by focusing on raising household income, or changing attitudes toward educating daughters.

3.6.2 Random Slopes Model

The random-intercepts model of section 3.6.1 assumed that the effects of each of the independent variables was fixed across communes, and across households within communes. In order to test the validity of this assumption, it is possible to estimate a *random slopes model*, that allows both the intercept and the coefficient (slope) of explanatory variables to vary randomly across higher level units. The particular focus here is on commune-level effects of household wealth on school attendance. The model takes the form

$$\ln\left[\frac{p_{ijk}}{1-p_{ijk}}\right] = \beta_0 + \beta_1 x_{1ijk} \ldots\ldots + \mu_{0k} + \mu_{1k} x_{1_k} + \varphi_{jk} + \varepsilon_{ijk} \qquad [12]$$

Column 5 of table 3.8 reports the results of equation [12], which allows the slope of mean-centered wealth to vary across communes. The effect of wealth on the log-odds of attending school in commune *k* (i.e. the average effect of wealth) is given by $\widehat{\beta_1} + \widehat{\mu_{1k}}$, which in this case is estimated to be 0.498 + $\widehat{\mu_{1k}}$. The between-commune variance in the effects of wealth, $\widehat{\mu_{1k}}$, is estimated to be 0.037.

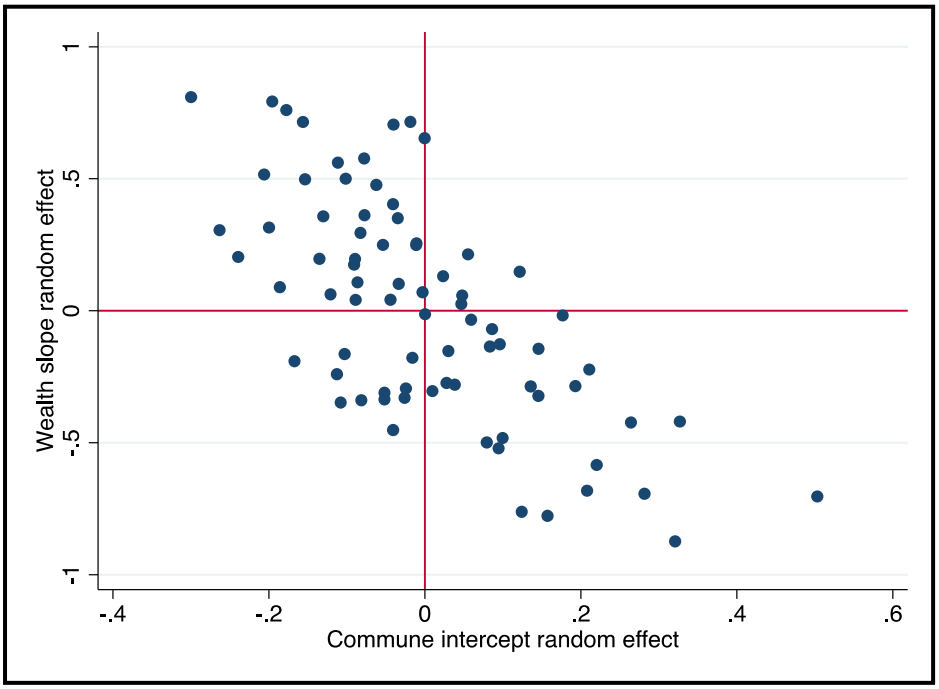
The between-commune variance in attendance, μ_k, falls from 0.241 to 0.214, which suggests that the distribution of wealth does indeed vary across communes (otherwise $\mu_k$would have remained unchanged). The estimated *commune intercept - wealth slope* covariance is negative (-0.043) which shows that those communes with below average primary school attendance rates (i.e. where $\mu_{0k} < 0$) also tend to have above average effects of wealth (i.e. where $\mu_{1k} > 0$). Figure 3.8 plots μ_{1k} against μ_{0k}.

In terms of potential policies to increase school attendance, this exposition is useful.

Communes in the upper left quadrant represent those where the effect of wealth on attendance is above average, whilst school attendance itself lies below average. Therefore, these communes represent those areas where the greatest improvements in school attendance rates could be realized through policy interventions that either raise income or lower the cost of schooling. The analysis of section 3.5 suggests that by lowering average distance to school (by building more schools, improving the road network etc.), this might free up time for boys to work (and as such contribute to family income) alongside studying, thus lowering the opportunity cost of attending school. This provides a clear example of how supply and demand side considerations work hand-in-hand to determine whether a child is sent to school.

Conversely, those communes lying in the lower left quadrant represent communes where attendance rates are below average, but the effect of wealth is also limited; increasing wealth levels here might have a limited impact on school attendance rates.

Figure 3.8 *Commune slopes vs commune intercepts*



3.7 Conclusion

This chapter has sought to shed light on the determinants of primary school attendance rates in Benin, a country that despite seeing almost unparalleled improvements in school attendance, gender parity and completion over the last two decades, has been practically ignored in the literature. The results presented here have presented numerous insights.

Firstly, the analysis of Beninese statistics has echoed the findings of, for example, Sandefur and Glassman (2015) who found that administrative statistics (in this case either from INSAE or UIS) overstated school enrolment compared to household surveys (DHS). For the school year 2005-06, DHS estimates of enrolment (attendance) were around 15 percentage points lower than those from UIS or INSAE. Secondly, large regional disparities existed during the year of study (2005-06) and indeed, still do in the most recent DHS data. Analysis at the commune level has been made possible due to detailed education statistics made available by INSAE for the years 2003-2010. The empirical analysis has presented strong evidence that factors on both the demand and supply side are predictors of whether or not a child will be enrolled in school. Richer households, those following a Christian religion and those with more educated household heads were more likely to send their children to school. Despite the narrowing gender gap in Benin at the time, girls still faced a lower likelihood of attending school than boys; adopted girls seemed at the greatest disadvantage. Focusing on the role of child labour, the empirical results highlight that those boys that worked alongside studying faced a higher opportunity cost of travel time to school. In particular, a larger distance to school for boys that worked in the field decreased the likelihood that they would attend school, to a much greater extent than those not working.

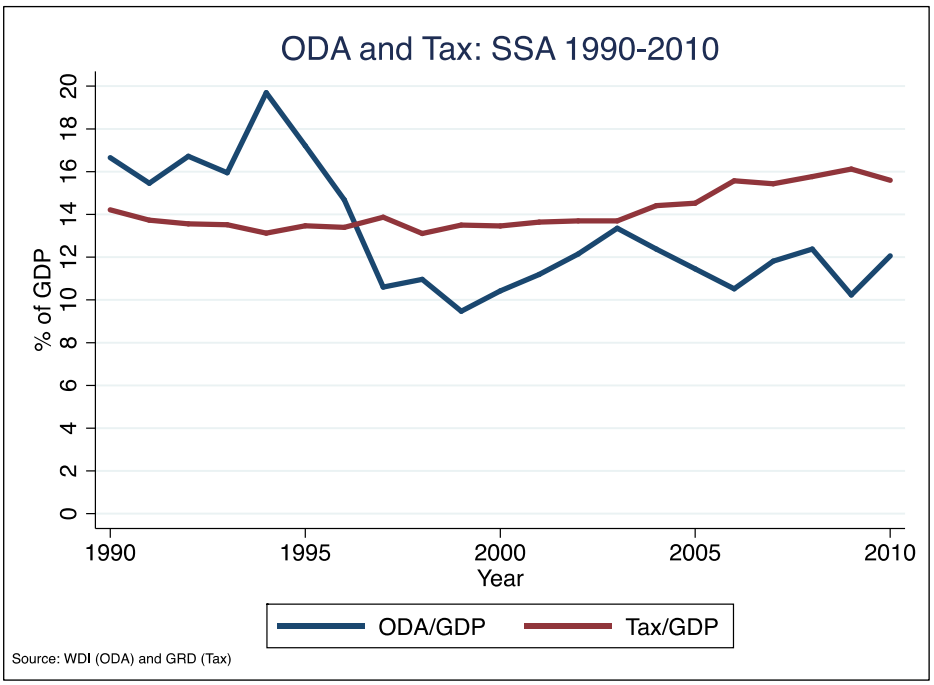
Whilst much of the literature investigating the determinants of school attendance or enrolment has acknowledged that factors at the community or state level might play a part in determining enrolment trends, many studies continue to ignore higher level clustering in the data. In an attempt to shed light on the large regional disparities that exist in Beninese primary school attendance, the present work has explicitly modelled higher level variation in school attendance by using a multilevel modelling strategy. A three-level random intercepts model, at the individual, household and commune level, highlighted that there were a number of communes where primary school attendance was significantly lower than average. After controlling for individual, household and commune-level factors, the model is able to explain a large portion of the between – commune variance in attendance rates. However, it also suggests that much of the residual variance in attendance between communes is actually due to factors at the household level. A random slopes model suggests a number of communes where average school attendance is below average and the effect of household wealth is above average. Such regions might stand to benefit most, in terms of school attendance rates, from government policy that raised household incomes, or reduced the costs of schooling.

Chapter 4: Tax Structures, Economic Growth and Development

4.1 Introduction

The past five decades have seen development economists pore over the relationship between foreign aid and economic growth. However, today the majority of low or middle income countries actually receive significantly more revenue from domestic tax receipts than from aid. Figure 4.1 illustrates that the average Sub-Saharan African country has collected more from domestic tax receipts than from official development assistance (ODA) for almost 20 years now. Indeed, a commitment to "…strengthen[ing] domestic resource mobilisation" has been embodied in the Sustainable Development Goals (UNSTATS, 2016:35) as part of a growing focus on increasing domestic resource collection in developing countries.[73]

Figure 4.1. *ODA and Tax receipts in Sub-Saharan Africa 1990-2010*



Yet, we know comparatively little about the effects of increased tax collection, or changes in tax structure, on economic growth and development. In part, this has been due to the paucity of data. Aid data, for example, is recorded by donor countries or multilateral organisations, whilst tax revenues are handled by individual countries' revenue authorities that often lack the administrative capacity to ensure that they are accurately recorded. The

[73] The UN has ratified two official indicators for Goal 17.1 (Strengthen Domestic Resource Mobilisation… …to improve domestic capacity for tax and other revenue collection). (i) total government revenue (% of GDP) and (ii) the proportion of domestic budget funded by domestic taxes.

primary source for researchers interested in tax has traditionally been the IMF's *Government Finance Statistics* (GFS), however a glance at these statistics show that they are of limited use for empirical analyses on developing countries due to, amongst other reasons, extensive missing data. Indeed, IMF researchers themselves have often seen fit to construct their own *ad hoc* datasets for empirical work (Prichard, 2016, provides examples). Crucially, these are often unavailable to researchers wishing to replicate or challenge their findings. One such study is that of Acosta-Ormaechea and Yoo (2012), which examined the relationship between tax structures and GDP growth rates, finding that revenue neutral shifts away from consumption and property taxes toward income taxes were harmful for growth in the long run. However, the study relies on a dataset that is not publically available.[74] Furthermore, the aforementioned work claims to present results for 'Low-' and 'Middle-Income' countries, but without explicitly naming these countries, it is nigh on impossible to draw any policy conclusions whatsoever. Worryingly, the results of studies such as this have been cited in IMF policy documents that discuss the impact of tax structures on economic growth.[75]

This study uses the ICTD-UNU WIDER Government Revenue Dataset (henceforth GRD) in order to extend, and also challenge, existing work on the effect of revenue neutral changes in tax structure on economic growth. In particular, the empirical estimations replicate the aforementioned work of Acosta-Ormaechea and Yoo (2012), before considering some extensions and robustness checks made possible due to the GRD, which presents a significant improvement in terms of availability of revenue data for developing countries, allowing for a range of new insights and policy-relevant analyses.[76,77] Specifically, we are able to gain an insight into the effects of trade liberalisation on economic growth in developing countries. The on-going trends of globalisation and IMF support for moves toward consumption taxes such as the value-added tax (VAT) have seen many developing countries' reliance on trade taxes decrease, measured as either a share of total tax or as a percentage of GDP. However, little is known about the impact of such

[74] When contacted, the authors were unwilling to share their dataset or their *Stata* .do files to assist with replication.
[75] See, for example, IMF (2011) or IMF (2015)
[76] In keeping with Clemens (2015), the empirical part is best described firstly as a *replication*, as I follow the same specification as the Acosta-Ormaechea and Yoo (2012) study. As the regressions contained herein contain a larger sample and consider some additional specifications, the work might also be described as an *extension* (which Clemens categorises as a kind of *robustness test*.) Unfortunately, it is not possible to classify the present work as a pure replication study as the authors of the original study were unwilling to share details on the countries included in their sample.
[77] Prichard (2016) provides an overview of the GRD, whilst Prichard *et al.* (2014) covers its construction in depth.

structural shifts in taxation. Baunsgaard and Keen (2010) highlight that revenue-recovery following the changes in many low-income countries has been poor; less still is known about the impacts on GDP growth rates. Acosta-Ormaechea and Yoo (2012) report that the majority of their findings did not hold for 'low income' countries, blaming the poor quality of data and crucially did not explore the effects of structural shifts away from trade taxes toward domestic consumption taxes.

The econometric analysis in this study uses the Pooled Mean Group (PMG) estimator (Pesaran, 1999) in order to estimate the effect of revenue-neutral changes in tax structure on economic growth rates. This primarily follows the approach taken in, for example, Arnold *et al.* (2011), Xing (2011) or Acosta-Ormaechea and Yoo (2012), but crucially extends the analysis to cover a number of developing countries. In a broad sense, the results find support for those in the aforementioned studies; revenue neutral (i.e. holding constant the tax ratio) shifts away from consumption or property taxes, toward income taxes, lead to lower long run GDP growth. However, no support is found for the findings of Arnold *et al.* (2011) that corporate income taxes (CIT) are the most harmful for growth. Nor is there broad support for the previously reported finding (Arnold *et al.* 2011; Xing, 2011; Acosta-Ormaechea, 2012) that revenue neutral increases in property taxes are good for economic growth. Turning to the effects of trade liberalisation (as measured here by revenue-neutral shifts away from taxes on international trade), the results suggest that for lower-middle income countries, there have been positive effects on GDP growth rates, but that for low income countries, the effect is insignificant or potentially negative. Indeed, at times the results differ dramatically between income groups which highlights that there is no *one size fits all* relationship between tax structure and growth – this is perhaps the key contribution of the empirical estimations presented herein. Subsequently, I conclude that no one policy prescription can be advised to all developing countries; certainly, any policy advice given to low income countries, using evidence from studies based on high income countries, would appear to be somewhat misguided.

The rest of this chapter is structured as follows. Section 4.2 provides an overview of the economic theory linking taxation and economic growth, before reviewing the related empirical literature. Section 4.2 also provides a brief discussion of why tax policy and the tax mix differs in developing countries. Section 4.3 introduces the data used here and examines the trends in tax structures for the sample. The following section, 4.4, outlines the empirical approach and results of the PMG estimations along with extensions,

robustness checks and limitations, are presented in section 4.5. Section 4.6 provides a discussion and conclusion.

4.2 Tax, Growth and development

4.2.1 In theory

This section provides a brief overview of the theoretical predictions surrounding the relationship between taxation and GDP growth.[78] Considering first the *tax ratio*, it is not immediately clear whether a higher level of taxes will have a positive or negative effect on output or growth. Whilst higher tax rates distort the incentive for individuals and firms to engage in activities that contribute towards higher levels of output, higher revenue provides the government with a greater ability to enact policies that can increase the productive capacity of the economy (e.g. as investment in education and training or subsidies for research and development). Thus as Arnold *et al*. (2011: F59) note, the relationship between the levels of taxation and output is more likely to be driven by 'societal choices as to the appropriate level of public spending'.

The neoclassical growth model (*inter alios* Swan 1956; Solow 1956) does not provide scope for assessing the potential for fiscal policy to affect the long-run steady-state growth rate; in this model a change in the tax rate may lead to a shift in the steady-state growth *path*, but not in its *slope*. In order to find theoretical predictions of the effects of fiscal policy on output growth however, one must look to works such as Barro (1990), King and Rebelo (1990) or Mendoza *et al*. (1997). Specifically, these models consider changes in the marginal tax *rate*. King and Rebelo's model (1990: 130) considers the effects of a rise in the rate of output tax '…applied equally to all sectoral activities'. The authors conclude that whilst 'taxation may affect the growth rate in a quantitatively important way … the magnitude of this influence depends … on the production and tax structure' (King and Rebelo 1990: 140).

The endogenous growth model outlined in Mendoza *et al*. (1997) however goes further by considering the effects of marginal tax rates on physical capital, human capital and consumption. The model predicts that whilst all three affect the 'net after-tax rate of return on physical capital', consumption taxes will do so only indirectly through the labour-leisure

[78] A comprehensive review can be found in Myles (2007).

choice, which in turn affects the ratio of capital to labour used in production. Higher consumption taxes increase the cost of consumer goods, in turn reducing the reward for working, thus impacting the labour supply (Arnold *et al*. 2011). Taxes on physical or human capital, however, influence growth through both direct effects on labour supply and indirect effects on the labour-leisure choice.[79] Whilst the magnitude of these impacts is dependent on factors such as the elasticity of labour supply, the predictions of the model are clear: there are fewer channels through which consumption taxes can distort growth (as measured by the return on physical capital) than with the other two tax categories. Thus the take away message from the relevant theoretical models is that taxes on physical and human capital (i.e. corporate or personal income taxes) might be more harmful to economic growth than consumption taxes.

4.2.2 Existing empirical work

Following Easterly and Rebelo's claim (1993:442) that "The evidence that tax rates matter for growth is disturbingly fragile", the empirical growth literature has made great strides towards more accurately isolating the effects of tax policy. It certainly appears at this point that, contrary to Easterly and Rebelo's claim, the evidence suggests that taxes do matter for growth; *which* taxes, *where* and *by how much* they affect growth are perhaps more stimulating, relevant, and due to the GRD, answerable questions.

Early studies examining the relationship between tax and growth simply included tax receipts as a share of GDP as a regressor. Plosser (1992) and Easterly and Rebelo (1993) both found a negative relationship between income taxes and growth. Yet both authors advise caution in interpreting the results as causal. The former points to the problem of collinearity among explanatory variables, whilst Easterly and Rebelo (1993) warned that their result was heavily dependent on the other covariates included. There are various shortcomings with using tax receipts' share of GDP as the fiscal variable of interest. Easterly and Rebelo (1993) recognise that this can only be interpreted as the marginal rate of, for example, income tax, if income taxes were proportional. Martinez-Mongay (2000) argues that the main problem with using this method is that GDP itself is determined by the income from labour and capital. Furthermore, he notes that disaggregating individual tax flows into their share of GDP is also a poor proxy as changes in these figures could be a result of either a change in the tax rate, or a change in the relevant tax base's share of GDP, which might not necessarily be a result of tax law.

[79] See Mendoza *et al*. (1997:102-106) for detailed derivations and results.

Kneller *et al*. (1999) claim that the majority of previous research ignores the fact that by focusing on just one side of the budget (i.e. on expenditure or taxation, but not both) results may be biased. Classifying taxation as 'distortionary' or 'non-distortionary', Kneller *et al.* (1999) predict that shifts towards the latter form of taxation (from distortionary) will have a positive effect on growth.[80] The premise of this classification lies in their claim that the 'most relevant distortions' come from the incentives to invest rather than the labour-leisure choice, which is the only facet affected by consumption taxes. Results suggest that distortionary taxes do indeed retard GDP growth. Estimates point to an increase in GDP growth of between 0.1 and 0.2 per cent following a 1 per cent decrease in distortionary taxation. These results hold after rigorous robustness checks in Bleaney *et al.* (2001). By and large, the result that increases in distortionary taxation, offset by decreases in non-distortionary revenue, are harmful to GDP growth, is in line with the theoretical predictions of Barro (1990), Barro and Sala-i-Martin (1995) or Mendoza *et al*. (1997).

More accurate estimates of the marginal tax rate, than can be provided by simply expressing revenue as a share of GDP, are difficult to achieve. In order to fully and accurately estimate the effects of tax rates on GDP, one would require information not only on each of the marginal rates but also on the income distribution, which, for a large sample of countries over many years, is difficult to obtain or accurately estimate. Attempts have however been made; Lee and Gordon (2005) use the top corporate tax rate, finding a negative and significant relationship with per capita GDP growth. Coefficient estimates point to around a 0.6 percentage point increase in growth arising from a 10 per cent cut in the top corporation tax rate. Mendoza *et al*. (1994) construct their own 'effective tax rates' for consumption, labour and capital, which are computed as the ratio of the difference between the pre- and post-tax value of (consumption, labour or capital) income to the value of these incomes at pre-tax prices.[81] Using the effective tax rates, Mendoza *et al*. (1997) test the theoretical model outlined in the same study, finding that a 10 percentage point decrease in labour (capital) income taxes leads to an increase in the investment rate of 1.8 (1.0) per cent. However, no significant relationship was uncovered when using GDP growth as the dependent variable. Whilst these methods may be a closer fit to the relevant theoretical models, in the sense that they attempt to estimate the marginal tax rate, the data

[80] They define distortionary taxes as taxes on income and profit, social security contributions, taxes on payroll and manpower and property taxes. Non-distortionary taxes are therefore taxes on consumption, including trade taxes.

[81] Lee and Gordon (2005) argue that the effective tax rates measure nothing more than an average tax on labour income, as opposed to the marginal rate.

requirements pose a severe limitation to applicability, especially across a large panel of countries. This perhaps goes some way to explaining why such an approach has not been regularly replicated in the empirical literature.

An innovative approach that has been taken in more recent studies is to examine *revenue-neutral* changes in the tax structure – that is, the effect on growth of changes in the tax structure, holding total tax receipts constant. By imposing the revenue-neutrality constraint, it is possible to 'avoid the difficulty of taking account of how any changes in aggregate revenue might be reflected in changes in public expenditure' (Arnold *et al.* 2011: F59). This approach is appealing in the sense that it allows the researcher to make use of datasets with rich information on the various components of taxation (i.e. the GRD) without requiring a similar level of coverage on the expenditure side. Arnold *et al.* (2011) find that a revenue-neutral increase in the share of income taxes (offset by a decrease in the share of consumption and property taxes) reduces GDP per capita in levels for a panel of OECD countries. Specifically, their results suggest that a one percentage point increase in the share of income tax revenue leads to a reduction in the long-term level of GDP of between 0.25 and 1 per cent. Moreover, they find that corporate income taxes have a stronger negative effect on GDP than do personal income taxes and that property taxes are most growth friendly. Thus the authors find support for the theoretical predictions that consumption taxes are less distortive to growth than personal or corporate income taxes. However, Xing (2011), using a similar sample, did not find evidence in support of the finding on the role of corporate income taxes. Acosta-Ormaechea and Yoo (2012) extend the analysis of Arnold *et al*. (2011) to a broader panel of countries, but consider per capita GDP growth as opposed to GDP in levels. Similar results are uncovered: revenue-neutral increases in income taxes, offset by decreases in consumption and property taxes, lead to slower long-term economic growth rates to the tune of between 0.07 and 0.14 per cent for a 1 percentage point shift.

Despite disparities over the variable used to proxy the marginal tax rate, it seems that results of studies in this field have reached similar conclusions. Kneller *et al*. (1999), Bleaney *et al*. (2001), Lee and Gordon (2005), Arnold *et al*. (2011) and Acosta-Ormaechea and Yoo (2012) all present evidence, using a variety of different methods, that income taxes are more harmful for GDP or GDP growth than are consumption taxes. Specifically, some of these papers separate the effects of personal income taxes (PIT) and corporate income taxes (CIT), arguing that the latter are the most harmful for growth.

Returning to the questions posed at the start of this section, the existing evidence does suggest that taxes on corporate and personal income distort growth to a greater extent that those on consumption. By *how much* these taxes affect growth is wholly dependent on the proxy used as the tax variable in each case; the estimates discussed above however suggest only a modest impact of changes in tax rates or structure on economic growth. Turning to the question of *where* these effects have been seen, the vast majority of the aforementioned studies are consistent in that, to date, they have only considered OECD countries. The only papers to consider the effects amongst non-OECD (and specifically developing) countries have been those by Lee and Gordon (2005) and Acosta-Ormaechea and Yoo (2012). Whilst the former paper makes no attempt to distinguish between high- and low-income countries, Acosta-Ormaechea and Yoo (2012) found that their main results held for high- and middle- but not for low-income countries. Crucially, no indication is provided of the exact sample of countries included. However, given that their data is compiled form the IMF's GFS, it is easy to see that the group of countries they call 'Low Income' might well, in truth, be middle-income.[82] Many previous studies have likely confined analyses to OECD countries as a result of (for example) data availability. However, the GRD presents a first opportunity to carry out similar analysis on a panel of developing countries.

4.2.3 Tax policy in developing countries

This section provides a short overview of the features of developing, or low income, countries that explain the significant differences in the administration of fiscal policy compared to developed, or high income countries. Broadly speaking, Thirlwall (2006) notes that there are two main aims for fiscal policy with regards the financing of development. Firstly, to maximise the savings capacity of the economy by keeping as close to full employment as possible and secondly, to design a tax policy that can raise the marginal propensity to save as much as possible without introducing distortions or disincentives to work.[83]

One crucial characteristic of fiscal policy in developing economies is the difference in tax structure and, in particular, the relatively heavy reliance on trade taxes such as import tariffs when compared to richer nations. Whilst this might be as a result of protectionist policies,

[82] There are few, if any, cases where the GFS contains a series of sufficient length for any low income country to carry out the analysis described in their paper. Furthermore, those countries labelled 'low-' and 'middle-' income are done so not according to, for example, the World Bank's income classification, but according to an *ad-hoc* procedure, based only on the income levels of those countries in their dataset.

[83] For a concise discussion of the 'sensitive' role that taxes must play in developing countries see, for example, Tanzi and Zee (2000).

it is also the case that border taxes are cheaper and easier to collect than income or consumption taxes: tariffs can be collected at ports or border crossings and thus the administration of such charges is easier organised and collected than that of say, an income tax on workers across the country. A number of works provide further insight into why tax structures differ: Riezman and Slemrod (1987) consider that the level of education amongst the workforce and population density influences the costliness of tax collection, arguing that collecting trade tariffs requires a much smaller number of educated civil servants compared to those required to collect an income or sales tax. The authors find the level of education and population density to be negatively correlated with trade tax's share of total revenue, adding weight to their claim that both of these variables are suitable proxies for estimating the cost of collecting taxes. The level of education is of course important on both sides: those actually paying the taxes must also understand the system and be able to keep track of accounts. This point is echoed by Tanzi and Zee (2000), who also note that unreliable communication infrastructures (mail, telephone *etc*.) can further hinder the efficiency of tax administration.

Thirlwall (2006) argues that income taxes might not be an effective means to raise revenue due to the fact that the majority of citizens might not actually earn enough to even be eligible to pay; he notes that only around 30 per cent of national income is taxable in developing countries, compared to some 70 per cent in developed nations. Yet too heavy a reliance on trade taxes has well-understood growth implications; the efficiency gains from exposure to international competition can greatly enhance the long run growth potential not only for specific industries, but the economy as a whole. Trade liberalisation does however place significant revenue constraints on many low-income countries, which might find limited scope to meet these needs by increasing standard consumption tax rates (IMF, 2011).

Another feature that helps to explain the tax structures of developing economies is the prevalence of the informal sector, from which tax revenues are not (usually) collected. According to the latest estimates from Schneider *et al*. (2010), the size of the average shadow economy in Sub-Saharan Africa was some 38.4% of GDP. This compares to an average of just 13.5% in high-income OECD countries.[84] Where a large amount of economic activity occurs in the informal sector, not only does this reduce the government's

[84] The country in SSA with the largest informal sector in 2007 was Zimbabwe, at 62.7%, whilst the smallest in the OECD was Switzerland, at 8.1%.

Potential tax revenues but it might also have longer-term implications for economic growth.[85] By remaining unregistered, firms forgo access to government public goods such as business courts, the chance to bid for government contracts, or access to formal credit markets that might enable them to expand production (Woodruff, 2013). Furthermore, Woodruff (2013) points out that by keeping their costs and revenues off the books (either through an unwillingness to pay taxes or an ineptitude with regards accounting procedures), small firms may not fully understand their own costs structures. This might result in either a loss of business in the case where goods are overpriced, or loss of revenues when under-priced. In this sense, remaining a part of the informal sector has important implications not only for individual firms, but also for the economicy as a whole.[86]

Yet it is often (prohibitively) costly to register a business in developing countries and as such, many informal sector firms are unlikely to do so unless coerced or forced. In Bolivia, for example, the cost of setting up a formal sector business can be as high as 260 per cent of GDP per capita, compared with an average of 34 per cent and a minimum of just 1.4 per cent in Canada (Djankov *et al.* (2002)). Alongside the fact that it is costly to administer taxes on income or profits (especially in light of the probable low profits made or incomes earned in small firms currently operating in the informal sector), this makes it unlikely that tax reform in developing countries would emphasise taxing the informal sector. The IMF (2011) strongly promotes the view that the greatest revenue improvements can be realised by tackling tax avoidance by high-earning individuals. This might, for example, take the form of legal loopholes, or the use of tax havens. It is clear that this problem is not specific to low-income countries, but better enforcement might lead to larger gains in such nations.

As a result of such issues in developing countries, much of the reform that has taken place in over recent decades has seen introductions of domestic consumption taxes such as the VAT, as opposed to increases in income or profit taxes. Indeed, this is certainly a change that has been encouraged by the IMF (2011), who describe the VAT as 'the standard policy prescription' for countries battling the revenue challenges that come from trade liberalisation. Specifically, they note that the VAT is able to raise revenue in a way that is

[85] Estimates for the size of the informal sector come from Schneider *et al.* (2010) and are only available for 1999-2007. Pooled OLS regressions provide further evidence that the size of the informal sector is negatively and significantly related to the level of tax receipts (in per cent of GDP). Results not shown.

[86] Of course, whilst the incomes and profits from the informal sector remain untaxed, such activity is, by definition, also not counted in National Income figures. As such the existence of a large shadow economy is important in our context (i.e. relating fiscal policies to changes in national income) only in the sense that it dictates that the tax mix will be more heavily reliant on those activities that can be taxed, such as international trade.

less distortive to economic activity than alternatives, whilst also having the advantage of being simpler to administer and comply with than, for example, an income tax.[87]

Tax structures in less developed countries are rather different from that of developed nations. Factors such as illiteracy, low population density, reliance on subsistence agriculture or the informal sector necessitate that policymakers are faced with very specific challenges in raising the revenue required to provide the necessary public goods that facilitate growth and development. Other issues (See IMF, 2011, for a full discussion) affecting tax collection might be the volatility of revenues from natural resource extraction (a particularly major concern in countries heavily reliant on exporting resources), or the damage caused to the tax base and administrations by conflict. Thus the succinct conclusion of Tanzi and Zee (2000:300) is quite fitting:

> "*In developing countries, tax policy is often the art of the possible rather than the pursuit of the optimal*."

4.3 Data & Trends

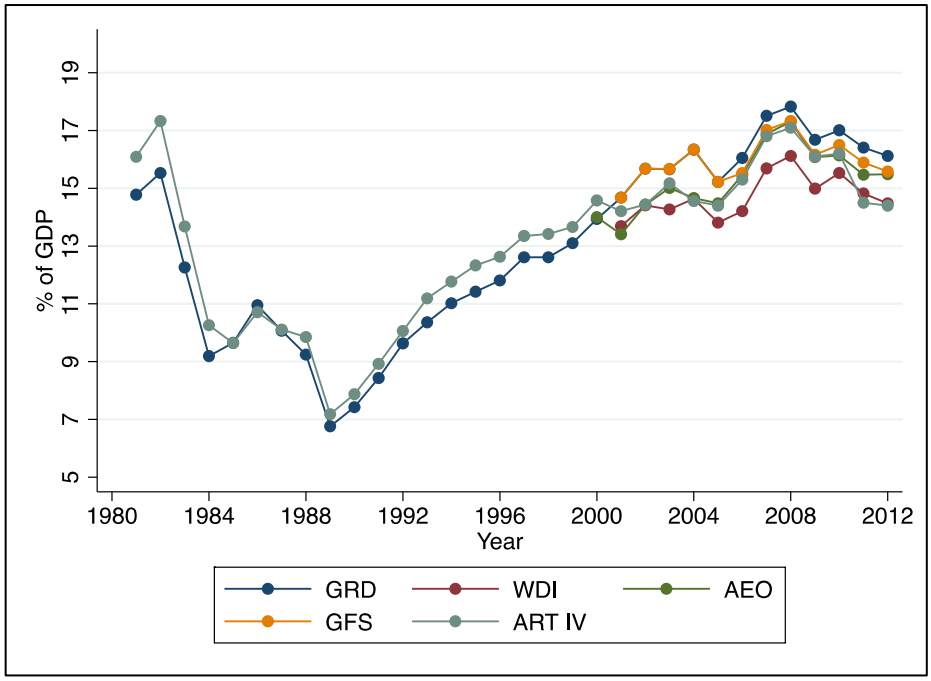
4.3.1 The Government Revenue Dataset

The tax revenue data for this chapter comes from the ICTD UNU-WIDER GRD. This dataset is the result of an effort to synthesise and combine data tax data from many sources, such as the IMF's *GFS*, OECD *Tax Statistics* and IMF *Article IV* Country Reports. Depending on the country in question, each of these sources report data to varying degrees of disaggregation and for various periods of time. Take the case of Benin, for example: whilst not an extreme example of the disparities across different sources, a look at the tax ratio by source is nonetheless interesting.

(Before discussing the data – an examination of the trend for Benin is insightful. The collapse of the socialist regime in the 1980s is accompanied with a large drop off of tax revenues, but these have been gradually improving since 1990 – in keeping with much of the continent.)

[87] Specifically, the IMF (2011) advocate that a *base-broadening* of the VAT would bring more significant revenue effects than increasing existing rates.

Figure 4.2 *Benin, Total Tax (% of GDP), by source*



Sources: ***GRD****: Government Revenue dataset.* ***WDI****: World Development Indicators.*
AEO*: African Economic Outlook.* ***GFS****: Government Finance Statistics*
ART IV*: IMF Article IV Country Reports.*

The line plots in figure 4.2 display the same series (total tax as a % of GDP) from 4 different sources, along with the GRD. There are a number of inherent issues with the picture presented here. Firstly, the most widely used source for revenue data, the GFS, only starts reporting tax revenue data for Benin in the year 2001. This is also the case for the WDI, whilst the AEO only has data from 2000 onward. Secondly, each of the sources report somewhat different figures. The disparity between sources can at times be close to 2% of GDP. Whilst this may not seem large, for a country only collecting government revenue amounting to around 15% of GDP, to under-measure by 2% is quite a sizable margin of error. One of the important reasons for this disparity, alongside the fact that different sources often report different revenue figures, is that they often use different underlying GDP figures. Indeed, a look at the underlying figures in local currency units (LCU) show that the WDI and GFS actually report identical figures for Benin, but when expressed as a percentage of GPD, the tax ratio differs wildly. The GRD, however, overcomes this problem by taking data from each source in LCU and dividing it by a common GDP series, namely the IMF's World Economic Outlook (WEO). Thus, in cases where different sources may report data for different time periods, it is often possible to merge these to form one series. For Benin, for example, the GRD uses data from IMF Art IV reports between 1980 and 1999, before switching to the GFS in the year 2000. The result is that for a large number of countries, the GRD is able to provide a long, consistent series which achieves large gains

in both analytical accuracy and completeness over any other one source. The GRD does not, however, make any attempt to alter the underlying numbers. The choice of source for each country depends on a number of factors but, generally speaking, the GRD will include revenue data that (i) can account for taxes accruing from resource extraction, (ii) provides a long and consistent series with high levels of disaggregation and (iii) extracts social contributions from tax. [88]

4.3.2 Examining Trends in the Data

The most recent release of the ICTD UNU-WIDER GRD contains some 6390 observations for 196 countries over the period 1980-2012/13. However, the econometric analysis here must rely on a smaller subsample of this data for a number of reasons. Firstly, it is crucial that a consistent time series is present for each country included. So, for example, if there was data for a country spanning 1980-1995 and 1996-2012, it would be dropped from the analysis. There are a number of countries where this is the case. The sample is also restricted to those countries with at least 20 years of consecutive observations, so that the *t* dimension is of sufficient length to carry out the PMG regressions. Thirdly, the GRD includes a number of 'flags' identifying potentially problematic data. Those observations flagged as '*Problem 1: Data not Credible*' or '*Treat With Caution: Data of Somewhat Questionable Quality'* are also excluded. Finally, the analysis is also restricted by the availability of some other covariates – specifically the measure of human capital (average years of schooling). The final sample for the baseline econometric analysis comprises 2657 observations for 100 countries.[89] The figures below, however, include more data where available, including for countries that are not included in the regression analysis.

Table 4.1 provides summary statistics for all variables included in the analysis for the sample as a whole, and also by income group (according to the latest available classifications from the World Bank). All tax variables come from the ICTD UNU-WIDER *GRD*. GDP growth is the change in (log) GDP per capita, taken from the World Bank's *World Development Indicators* (WDI). *Physical Capital*, also from the WDI, is (log) fixed capital formation expressed as a share of GDP. The variable *Human Capital* is the average

[88] A full account of the construction of the GRD can be found in Prichard *et al*. (2014).
[89] A list of these countries, by income group, is included in Appendix E.

Table 4.1 *Summary Statistics*

Variable		Obsv.	Avg.	Std. Dev.	Min.	Max.
Per capita GDP Growth	*All*	2753	0.018	0.042	-0.649	0.316
	HI	1236	0.020	0.033	-0.149	0.171
	UMI	722	0.018	0.046	-0.216	0.168
	LMI	392	0.022	0.031	-0.156	0.122
	LI	402	0.009	0.062	-0.649	0.316
Physical Capital	*All*	2753	0.216	0.067	0.000	0.748
	HI	1236	0.227	0.047	0.085	0.463
	UMI	722	0.222	0.069	0.000	0.484
	LMI	392	0.227	0.089	0.055	0.748
	LI	402	0.162	0.063	0.000	0.343
Human Capital	*All*	2753	7.510	2.988	0.796	13.247
	HI	1237	9.734	1.745	4.803	13.247
	UMI	722	7.186	1.824	2.836	11.289
	LMI	392	5.605	2.257	1.571	11.147
	LI	402	3.104	1.581	0.797	7.865
Population Growth Rate	*All*	2753	0.018	0.014	-0.033	0.139
	HI	1237	0.010	0.012	-0.025	0.111
	UMI	722	.023	0.012	-0.021	0.139
	LMI	392	0.023	0.009	-0.012	0.048
	LI	402	0.028	0.013	-0.033	0.069
Tax/GDP	*All*	2675	0.224	0.114	0.006	0.628
	HI	1236	0.301	0.096	0.011	0.506
	UMI	722	0.181	0.079	0.006	0.464
	LMI	392	0.173	0.105	0.047	0.628
	LI	402	0.11	0.045	0.018	0.326
Income Tax (share of total tax)	*All*	2752	0.444	0.177	0.000	0.894
	HI	1236	0.576	0.120	0.000	0.894
	UMI	722	0.387	0.147	0.128	0.799
	LMI	392	0.313	0.125	0.088	0.685
	LI	402	0.270	0.100	0.044	0.635
Personal Income Tax (Includes Payroll) (share of total tax)	*All*	2162	0.205	0.124	0.000	0.787
	HI	1117	0.266	0.126	0.000	0.787
	UMI	493	0.130	0.091	0.000	0.440
	LMI	288	0.142	0.068	0.000	0.412
	LI	264	0.156	0.088	0.004	0.518

Variable		Obsv.	Avg.	Std. Dev.	Min.	Max.
Social Contributions (share of total tax)	*All*	2499	0.137	0.143	0.000	0.476
	HI	1225	0.211	0.141	0.000	0.476
	UMI	722	0.104	0.117	0.000	0.413
	LMI	288	0.035	0.084	0.000	0.391
	LI	264	0.003	0.009	0.000	0.036
Corporate Income Tax (share of total tax)	*All*	2162	0.111	0.068	0.000	0.516
	HI	1117	0.092	0.059	0.000	0.484
	UMI	493	0.139	0.079	0.000	0.516
	LMI	288	0.133	0.069	0.030	0.391
	LI	264	0.112	0.051	0.017	0.258
Consumption & Property Tax (share of total tax)	*All*	2752	0.555	0.176	0.106	1.000
	HI	1236	0.424	0.12	0.106	1.000
	UMI	722	0.613	0.146	0.201	0.872
	LMI	392	0.687	0.125	0.315	0.912
	LI	402	0.73	0.099	0.365	0.956
Consumption Tax (Goods & Services + Trade) (share of total tax)	*All*	2690	0.528	0.194	0.100	1.000
	HI	1198	0.372	0.125	0.100	1.000
	UMI	702	0.599	0.152	0.157	0.870
	LMI	391	0.681	0.126	0.299	0.911
	LI	399	0.721	0.097	0.365	0.953
Tax on Goods and Services (Includes 'Other Tax') (share of total tax)	*All*	2690	0.374	0.14	0.000	0.870
	HI	1198	0.33	0.108	0.028	0.644
	UMI	702	0.4	0.163	0.017	0.870
	LMI	391	0.423	0.134	0.097	0.681
	LI	399	0.412	0.151	0.000	0.723
Trade Tax (share of total tax)	*All*	2691	0.154	0.176	0.000	0.888
	HI	1199	0.043	0.092	0.000	0.888
	UMI	702	0.199	0.168	0.000	0.700
	LMI	391	0.258	0.171	0.016	0.799
	LI	399	0.308	0.171	0.030	0.847
Property Tax (share of total tax)	*All*	2745	0.03	0.035	0.000	0.176
	HI	1236	0.052	0.036	0.000	0.176
	UMI	715	0.018	0.022	0.000	0.086
	LMI	392	0.006	0.013	0.000	0.072
	LI	402	0.009	0.017	0.000	0.094

years of schooling from Barro & Lee (2013).[90] Population growth is the growth rate of the working age (15-64) population, calculated from the *WDI*. The average GDP growth rate for the whole sample is 1.8%. Investment in physical capital is on average 22% of GDP. The average years of schooling across the sample is 7.5 years, however this ranges from less than one to over 13 years. Average years of schooling in low income countries is just over 3, increasing to almost 10 years in the average high income country. The tax ratio is, on average, 22%. The trends in tax mix, and how these differ between income groups, are outlined in figure 4.3.

Figure 4.3 *Tax Ratio & Structure by Income Group*

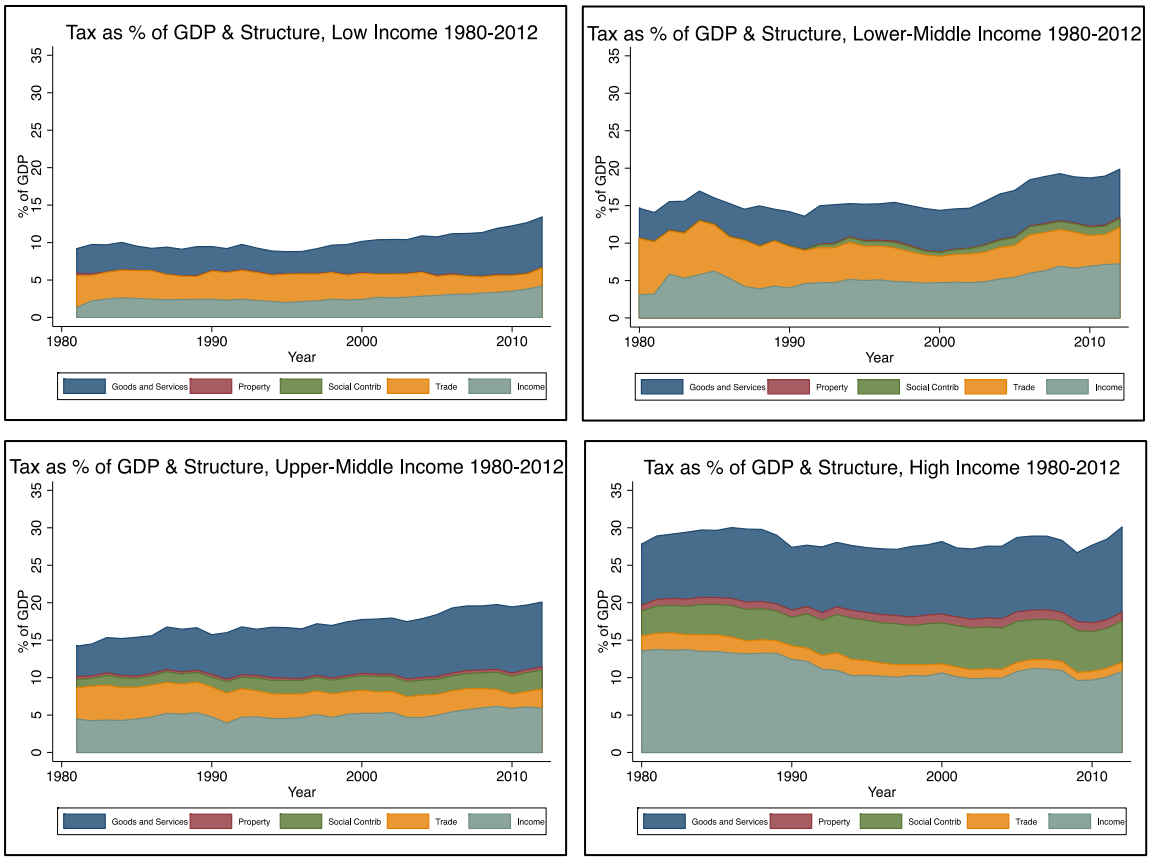


Figure 4.3 presents the average tax ratio and tax structure for each of the four World Bank income groups from 1980/81-2012. The tax subcategories shown are income (including personal income, corporate income and taxes on payroll and workforce), taxes on goods and services (including all domestic consumption taxes such as VAT, sales tax or excises; for simplicity, 'Other' taxes have been included in this category), trade, property and social contributions.

[90] Given that this data is only available at 5-year intervals, I use linear interpolation in order to gain a complete time series. This follows the approach taken in, for example, the Penn World Tables, whose human capital variable is highly correlated with the one used here (corr = 0.98).

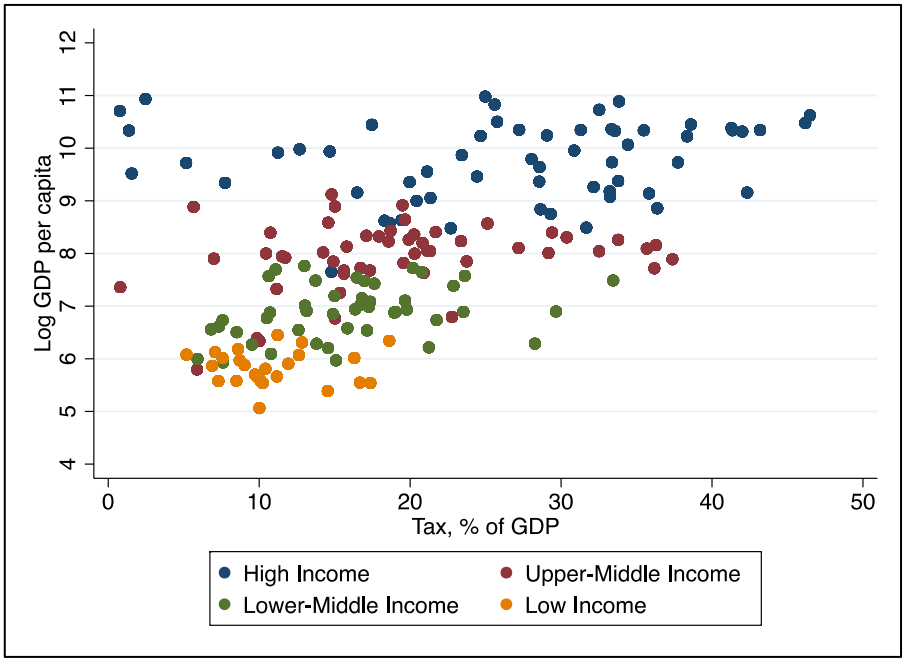
The tax ratio in high income countries stands at between 25 and 30 % for the period in question; this is dramatically more than in low (10 – 15%) or middle income countries (15-20%). The decade 2000-2010 has seen a notable increase in the tax ratio in low and middle income countries, whereas those classed as high income have seen their tax ratios remain fairly constant. The effects of the recent financial crisis are evident for high income countries in the late 2000s, where the average tax ratio dips by 2-3 percent, although no such effect is seen in low or middle income countries.

Turning to the tax structure, immediately clear is the initially high reliance on taxes on international trade in low and middle income countries. Over the last three decades, this has declined, largely to be replaced by taxes on domestic goods and services and to an extent, income taxes. The reliance on trade taxes in low income countries has more than halved, from around 40% of total tax receipts in 1980 to under 20% in 2012. There is a similar, if not so dramatic shift away from trade taxes in lower- and upper-middle income countries for the same time period. These trends undoubtedly reflect the on-going removal of many trade barriers and also the implementation of taxes on domestic consumption such as the VAT. The figure confirms that not only have trade taxes been falling in relative terms (i.e. as a share of total tax), but also in absolute terms (i.e. as a share of GDP).

Figure 4.3 also highlights an increasing reliance on income taxes and social contributions in those countries classed as lower-middle income across the same period. High-income countries, perhaps as a result of more efficient tax collection and administration capabilities, are typically more reliant on income taxes or social contributions. The relative stability of the tax structure in high-income countries over this period is also noticeable. Many are members of free trade areas such as the European Union (EU) and as such the replacement of trade barriers with consumption taxes may well have occurred before the timeframe in question here. It is also clear that only high income countries collect a significant amount of revenue from property taxation, averaging around 4-5% of total revenue for the period in question.[91]

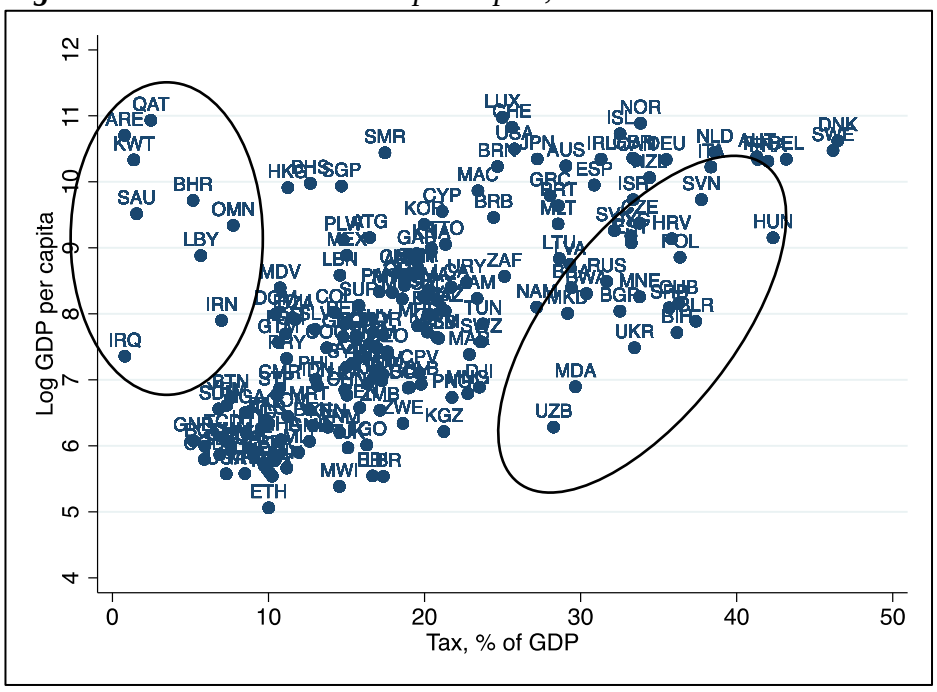
[91] NB. The data for the majority of lower-middle or low income countries comes from IMF Article IV Staff Reports. Frustratingly, these vary in the level of disaggregation reported. Where information on property taxes was not included (i.e. either there were no property taxes either collected or recorder), this assumed zero for simplicity.

Figure 4.4 *Tax Ratio and GDP per capita*



Turning to the relationship between tax and GDP per capita, figure 4.4 plots the average tax ratio against average log income per capita over the period 1980-2012. A few things are noticeable. Firstly, there is, overall a positive relationship between per capita income levels and the tax ratio. Secondly, the World Bank's income groupings can roughly predict the level of a country's tax ratio. i.e. the average tax ratio appears to increase with income group, but there are a number of outliers. A closer look at the outlying countries in figure 4.5 is interesting.

Figure 4.5 *Tax Ratio and GDP per capita, labelled*



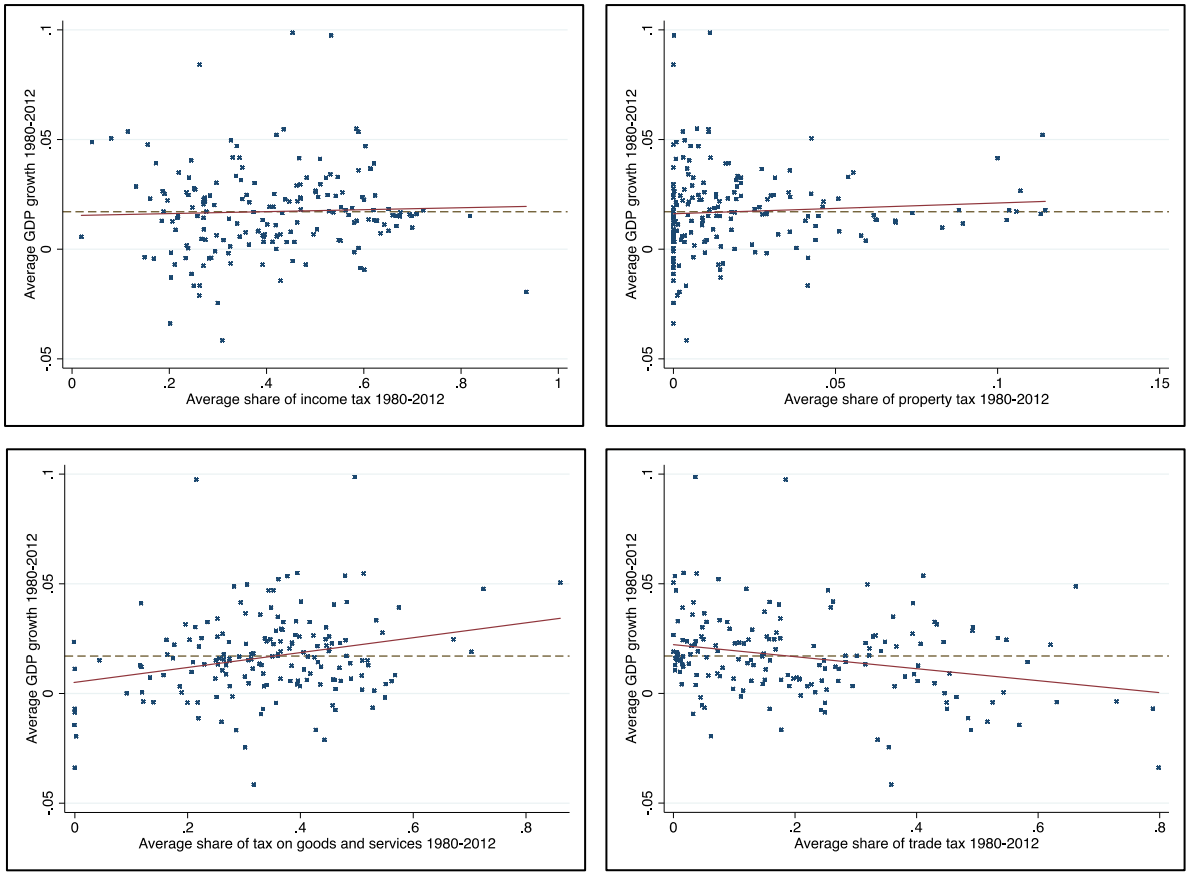
Those countries circled with very low tax ratios, but high GDP growth are exclusively oil-producing countries. Four others stand out as outliers, namely San Marino, Singapore, Hong Kong and the Bahamas. At the other end of the scale are almost exclusively former Soviet countries. Those circled, including Moldova, Ukraine, Belarus, Bosnia and Herzegovina, Hungary, Poland, Croatia, Montenegro, Serbia, Macedonia and Bulgaria all have tax ratios around 30-40 percent of GDP, but relatively low per capita income levels. Interestingly, this latter group of countries all have relatively high levels of social contributions (> 10% of GDP; performing the same analysis with taxes exclusive of social contributions entirely eliminates these outlying countries).

These figures are informative as they underline the fact that, for most countries, higher per capita GDP is associated with a higher tax ratio. The improved coverage in the GRD has allowed for a more complete picture of this relationship, with data points for around 185 countries included in the scatter plots.[92] However, the econometric analysis below is concerned with the effects of tax *structure* on GDP *growth*. Figure 4.6 presents a first look at the relationship between average GDP growth over the period 1980-2012 and the average share of tax that countries collect from income, property, goods and services and trade. The green dashed line shows the average GDP growth rate for the period across countries; the red line is a line of best fit.

There appears to be only a modest positive association between the share of taxes collected as income or property tax and the average GDP growth rate. However, the average share of taxes on goods and services appears to be positively associated with GDP growth rates, whilst the opposite is true for the average share of trade tax.

[92] There are not 33 years of data (1980-2012) for every country in the sample. Indeed, data for, for example the former Soviet states, only appears from the early 1990s. However, the patterns displayed in figures 2 and 3 remain if the sample is restricted to average tax and average income levels from 1991-2012 for all countries (not shown).

Figure 4.6 *Average GDP growth and Tax Structure variables*



4.4 Econometric Model

This section outlines the econometric model used to investigate the relationship between tax structure and economic growth. The approach follows that of, *inter alios* Arnold *et al*. (2011), Xing (2011) and Acosta-Ormaechea and Yoo (2012), by considering revenue-neutral changes in the tax structure, whilst controlling for the overall level of tax as a proportion of GDP. Imposing the revenue neutrality constraint allows for the consideration of the effects of changes in tax policy on growth, without the need to take account for how changes in tax policy might result from changes in government expenditure (Arnold *et al.*, 2011). Thus it avoids the requirement of similar levels of data on public expenditure as on tax revenue. Furthermore, given limitations on the availability of data on tax rates, this approach represents the best available proxy for the marginal tax rate which is considered in the relevant theoretical models (e.g. Mendoza *et al.*, 1997).

The empirical model estimated is an Error Correction Model (ECM) taking the form

$$\Delta g_{i,t} = -\phi(g_{i,t-1} + \alpha_1 I_{i,t-1} + \alpha_2 h_{i,t-1} + \alpha_3 n_{i,t-1} + \alpha_4 T_{i,t-1} + \sum \alpha_j \, TC_{i,t-1})$$

$$+\beta_{1,i}\Delta I_{i,t} + \beta_{2,i}\Delta h_{i,t} + \beta_{3,i}\Delta n_{i,t} + \beta_{4,i}\Delta T_{i,t} + \sum \beta_j \Delta TC_{i,t} + \tau_i t + \varepsilon_{i,t} \qquad [13]$$

where g is the growth rate of GDP per capita, I is the investment ratio (as measured by the share of fixed capital formation in GDP), h is a measure of human capital (average years of schooling), n is the growth rate of the working age population, T is the tax ratio and TC is a vector of tax composition variables, namely $n\text{-}1$ shares of different taxes in total tax. τ is a set of time dummies and ε is the error term.

The equation, estimated by Pooled Mean Group (Pesaran *et al.*, 1999) allows for simultaneous estimation of the long run coefficients and short run dynamics. The PMG estimator constrains long run coefficients to be equal across groups (countries), but allows short run coefficients and error variances to vary between groups.[93] The validity of this assumption is tested in section 6.1 below.

4.5 Benchmark Results

The full sample of 100 countries is included in table 4.2. The sample is restricted to countries where there are at least 18 consecutive years of observations for all variables, although on average, $t = 27$. All regressions contain short run dynamics as in [13], but only the long run coefficients are shown. The revenue neutrality constraint means that the interpretation of the coefficient on the included tax category's (categories') share in total tax is as follows: A percentage point increase in the share of tax revenue for the included tax component (components), implies a percentage point reduction in the share of total revenue from the omitted component (components). A summary of how taxes are subdivided is provided in Appendix F. In columns 1 and 2, the omitted category is the share of consumption and property taxes. The results suggest that revenue-neutral (RN) shifts away from consumption and property taxes toward income taxes have a negative effect on the long run GDP growth rate.

[93] The estimations are carried out using the *xtpmg* command in Stata (Blackburne and Frank, 2007). Newton-Raphson iteration is used.

Table 4.2 *PMG Results, Full Sample*

Dependent Variable: Δ(log) GDP per capita.

	1	2	3	4	5
Physical Capital	0.013	0.009	0.058***	0.047**	0.058***
	(0.019)	(0.023)	(0.020)	(0.020)	(0.020)
Human Capital	-0.030***	-0.051***	-0.033***	-0.029***	-0.033***
	(0.009)	(0.012)	(0.008)	(0.009)	(0.008)
Population Growth	-0.423***	-0.444***	-0.504***	-0.568***	-0.504***
	(0.133)	(0.161)	(0.152)	(0.143)	(0.152)
Tax / GDP	0.061**	0.096***	0.124***	0.092***	0.124***
	(0.029)	(0.008)	(0.032)	(0.029)	(0.032)
Tax Structure Variables					
Income	-0.062***			-0.047**	-0.034
	(0.018)			(0.021)	(0.022)
PIT		-0.068**			
		(0.030)			
Social Contribution		-0.090***			
		(0.033)			
CIT		0.045			
		(0.031)			
Consumption & Property					
Goods & Services			0.069***		0.036***
			(0.021)		(0.012)
Trade			0.034		
			(0.022)		
Property			0.058	-0.015	0.024
			(0.076)	(0.077)	(0.076)
Omitted Tax Variable	**Consumption & Property**		**Income**	**Consumption**	**Trade**
Observations	2657	1802	2290	2296	2290
Groups	100	68	84	84	84
Stationarity:	I(0)	I(0)	I(0)	I(0)	I(0)
5-Year Dummies	Yes	Yes	Yes	Yes	Yes

*Standard Errors in Parentheses. *P<0.1; **P<0.05; ***P<0.01*

Specifically, the coefficient estimate suggests that for a percentage point increase in income taxes, offset by a percentage point reduction in consumption and property taxes, the long run GDP growth rate decreases by 0.062 percentage points. Column 2 disaggregates income taxes into personal, social contributions and corporate. Here, the results suggest that a RN shift from consumption and property taxes toward personal income taxes or social contributions reduces long run GDP growth rates by 0.068 and 0.09 percentage points respectively. These findings echo those of Acosta-Ormaechea and Yoo (2012) although, notably, the coefficient estimates here are around 50% lower in magnitude than those reported in their study. In column 3, income taxes are omitted. It appears that RN shifts away from income and toward domestic taxes on goods and services have positive effects on long run GDP growth rates, but RN shifts toward trade or property taxes have no significant effects. In column 4, the omitted category is set to consumption taxes (goods and services + trade) and the results again suggest that, controlling for the share of income taxes in total tax, shifts away from consumption toward property taxes have no statistically significant effect on long growth rates. This result on the role of property taxes is in direct contrast to that of Acosta-Ormaechea and Yoo (2012) and Arnold *et al.* (2011), who found strong positive impacts of RN shifts toward property taxes on growth and income respectively.

Column 5 omits trade taxes, seeking to elicit the effect of a RN shift away from trade toward either income, property or domestic consumption taxes. This coefficient of 0.036 on domestic goods and services suggests that RN shifts away from trade toward domestic consumption taxes have had modestly positive effects on GDP growth rates. This result is intriguing, especially in light of the patterns in tax structure observed and discussed above, that showed that for many countries, shifts away from trade toward domestic consumption taxes have been the major structural change over the last 30 years. All specifications were tested for residual nonstationarity, by an Augmented Dickey Fuller (ADF) test.[94] As shown in table 2 , for all specifications the null of nonstationary residuals was rejected at the 5% level; in the majority of cases the null was also rejected at the 1% level.

In table 4.3, the same specifications are run, but the sample is separated into high, upper-middle, lower-middle and low income according to the World Bank's 2016 income classification. Column 1 shows that there are statistically significant negative effects on growth rates form RN shifts toward income taxes in high-income countries – to the tune of around 0.1 percentage points for a 1 percentage point increase. Disaggregating into PIT, Social Contributions and CIT, it actually appears that RN shifts away from consumption and property, toward corporate income taxes, have positive effects on long run growth rates in high income countries. This result conflicts with both theory and some existing empirical evidence. Columns 3 and 4 suggest that there are positive effects on long run GDP growth rates of RN increases in property taxes: a RN percentage point shift away from either consumption or income taxes leads to an increase in GDP growth rates of around 0.3%. This result affirms that presented in Arnold *et al.* (2011) and Acosta-Ormaechea and Yoo (2012), which found positive effects of RN shifts toward property taxes in OECD and 'high income' countries respectively. Whilst a 0.3% increase in long run GDP growth rates might sound quite high, it is worth noting that a revenue-neutral shift toward property taxes of 1 percentage point would be extreme in any one year: the average change in property tax's share of total tax for high income countries was just 0.02% of total tax revenue. There are no significant effects of RN shifts from trade taxes toward taxes on goods and services in high income countries. However, this result is not surprising; figure 2 highlighted that there has been little change in the share of taxes coming from trade toward goods and services in this group of high income countries.

[94] This test was carried out using the *pescadf* routine in Stata (Lewandowski, 2007). Output not shown.

Turning to upper-middle income countries, the estimates in column 6 clearly suggest that a percentage point shift toward income taxes away from consumption taxes is harmful for long run growth rates, to the tune of around 0.15 percentage points. Disaggregating income taxes (column 7) shows that the negative effects from social contributions and personal income taxes are again stronger than those from corporate income taxes. Interestingly, shifts toward property taxes from either income or consumption taxes also appear to have negative effects on GDP growth rates. Column 10 also suggests that RN shifts away from trade toward income taxes are harmful for growth, but those toward taxes on domestic goods and services are neither positively or negatively related to growth rates. Interestingly, it appears that RN increases in property taxes are actually *harmful* for economic growth in upper-middle income countries.

Columns 11-15 display the results for lower-middle income countries.[95] Intriguingly, RN shifts in tax structure toward income taxes do not appear to have any significant negative effects on long run GDP growth in this subsample; indeed the results in column 15 suggest that RN increases in income taxes offset by reductions in trade taxes might have positive effects on long run growth rates. The results in columns 13-15 suggest that RN shifts away from either income or consumption taxes, toward property taxes, again have negative effects on GDP growth rates. Specifically, the coefficient estimate points to around a 0.8-1.1% decrease in long run GDP growth rates for a percentage point increase in the share of taxes coming from property taxes. Column 15 shows that RN shifts toward domestic goods and services, offset by decreases in trade taxes, have positive effects on long run economic growth. Looking at the other country income groups, it would seem that this group of *lower-middle income* was driving the result in column 5 of table 4.2.

[95] It was not possible to obtain results for the disaggregated income tax categories for low and lower-middle income countries. The majority of data in the GRD for these countries comes from IMF Article IV Country Reports, where the level of disaggregation reported can fluctuate wildly between countries and over time for the same country. Often there is only one 'Income Tax' figure reported.

Table 4.3 *PMG results, by income group*

	High Income					Upper-Middle Income				
	1	2	3	4	5	6	7	8	9	10
Physical Capital	-0.082**	-0.036	0.017	-0.021	0.017	-0.141**	-0.158***	-0.091**	-0.136***	-0.091**
	(0.038)	(0.037)	(0.038)	(0.039)	(0.038)	(0.039)	(0.049)	(0.036)	(0.039)	(0.036)
Human Capital	-0.140***	-0.161***	-0.100***	-0.122***	-0.100	0.640**	0.030	0.067***	0.072***	0.067***
	(0.019)	(0.019)	(0.020)	(0.020)	(0.020)	(0.025)	(0.030)	(0.025)	(0.025)	(0.025)
Population Growth	-0.632***	-0.888***	-0.830***	-0.820***	-0.830***	1.573***	2.158***	2.049***	1.688***	2.049***
	(0.190)	(0.187)	(0.212)	(0.214)	(0.212)	(0.265)	(0.286)	(0.220)	(0.255)	(0.220)
Tax / GDP	0.213**	0.198***	0.195***	0.203***	0.195***	0.142**	0.122*	0.239***	0.180***	0.239***
	(0.048)	(0.049)	(0.047)	(0.048)	(0.047)	(0.056)	(0.072)	(0.054)	(0.059)	(0.054)
Tax Structure Variables										
Income	-0.101***			-0.024	-0.122*	-0.154***			-0.170***	-0.142***
	(0.035)			(0.037)	(0.068)	(0.033)			(0.034)	(0.036)
PIT		-0.113**					-0.348***			
		(0.048)					(0.079)			
Social Contributions		-0.120**					-0.303***			
		(0.47)					(0.064)			
CIT		0.163***					-0.056			
		(0.052)					(0.056)			
Consumption & Property										
Goods & Services			0.023		-0.099			0.120***		-0.022
			(0.036)		(0.064)			(0.036)		(0.034)
Trade Taxes			0.122*					0.142***		
			(0.068)					(0.036)		
Property Taxes			0.293***	0.327***	0.171			-0.345*	-0.630***	-0.487***
			(0.095)	(0.100)	(0.112)			(0.184)	(0.184)	(0.181)
Omitted Tax Variable	Consumption & Property		Income	Consumption	Trade	Consumption & Property		Income	Consumption	Trade
Observations	1202	1033	1121	1123	1121	695	377	663	663	663
Groups	42	36	38	38	38	27	16	26	26	26
5-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Table 4.3 *(continued)*

	Lower-Middle Income					Low Income				
	11	12	13	14	15	16	17	18	19	20
Physical Capital	0.117***	-	-0.016	0.115***	-0.016	0.152***	-	0.213***	0.132***	0.213***
	(0.027)	-	(0.016)	(0.025)	(0.016)	(0.056)	-	(0.064)	(0.057)	(0.064)
Human Capital	-0.039***	-	-0.004	-0.035	-0.004	0.044*	-	0.043	0.047***	0.043
	(0.012)	-	(0.006)	(0.010)	(0.006)	(0.025)	-	(0.028)	(0.025)	(0.028)
Population Growth	-1.133***	-	0.421***	-1.091***	0.421***	0.049	-	0.421	0.298	0.421
	(0.295)	-	(0.101)	(0.272)	(0.101)	(0.303)	-	(0.346)	(0.320)	(0.346)
Tax / GDP	-0.024	-	0.071***	0.002	0.071***	-0.241*	-	-0.313**	-0.272**	-0.313**
	(0.041)	-	(0.020)	(0.039)	(0.020)	(0.132)	-	(0.142)	(0.130)	(0.142)
Tax Structure Variables										
Income	0.001			0.016	0.081***	-0.120**			-0.173***	-0.193***
	(0.030)			(0.030)	(0.023)	(0.052)			(0.060)	(0.064)
PIT		-					-			
		-					-			
Social Contributions		-					-			
		-					-			
CIT		-					-			
		-					-			
Consumption & Property										
Goods & Services			-0.041		0.040***			0.205***		0.011
			(0.026)		(0.011)			(0.068)		(0.026)
Trade			-0.081***					0.193***		
			(0.023)					(0.064)		
Property			-1.109***	-0.822***	-1.028***			-0.271	-0.490**	-0.465*
			(0.278)	(0.242)	(0.274)			(0.226)	(0.235)	(0.247)
Omitted Tax Variable	Consumption & Property		Income	Consumption	Trade	Consumption & Property		Income	Consumption	Trade
Observations	374	-	359	360	359	386	-	383	383	383
Groups	16	-	15	15	15	15	-	15	15	15
5-Year Dummies	Yes	-	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Results for *low income* countries are shown in columns 16-20. The results here suggest that RN shifts away from consumption and toward income taxes again lead to lower long run GDP growth rates. RN shifts away from trade taxes, toward taxes on goods and services appear to have no statistically significant positive effect on GDP growth rates. This is, again, intriguing, considering that the data and graphs presented above suggest that the largest structural shifts away from trade toward taxes on goods and services occurred in those countries classed as low income. Revenue – neutral increases in property taxes, offset by decreases elsewhere, again appear to have negative effects on long run GDP growth rates.

4.5.1 Formal tests of differences in coefficients between subsamples.

The regression results above in Table 4.3 have shown that for different subsamples, the coefficients on key tax share variables have different magnitudes and occasionally different signs, depending on income group. This section presents a set of simple Wald tests to examine whether or not the differences between coefficients across income groups are statistically significant. It is, unfortunately, not possible to carry out this analysis using the PMG procedure as above; the approach followed involves adding dummy variables for each income group and interaction terms with the parameters of interest – dummy variables are, by nature, omitted from any panel data analyses and to include the additional parameters into a PMG specification causes the estimator to fail to converge. Thus I turn to a simple OLS regression.[96]

Table 4.4 replicates table 4.2, including dummy variables for three of the four income groups and interactions with the tax structure variables of interest.

A simple Wald test of the following null hypothesis is carried out:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 \qquad [14]$$

where $\beta_1 \ldots \ldots \beta_4$ refers to the coefficient on the interaction between the dummy and tax share variables for low, lower middle, upper middle and high income groups respectively.

[96] It should be noted that when the regressions presented above are run via OLS, the majority of coefficients have the same sign and are of similar magnitude.

Table 4.4 *Replication of Table 4.3, including tests for significant differences between estimated tax structure coefficients across income groups.*

Dependent Variable: Δ(log) GDP per capita. SR dynamics and time controls included but not shown.

	1	2	3	4	5
Physical Capital	0.106***	0.075***	0.109***	0.110***	0.109***
	(0.015)	(0.018)	(0.015)	(0.016)	(0.015)
Human Capital	0.009***	0.012***	0.010***	0.011***	0.010***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Population Growth	-0.213	-0.166	-0.172	-0.233*	-0.172
	(0.154)	(0.229)	(0.156)	(0.138)	(0.156)
Tax / GDP	-0.007	0.015	0.001	-0.006	0.001
	(0.013)	(0.016)	(0.014)	(0.013)	(0.014)
Income	-0.038***			-0.040***	-0.029
	(0.010)			(0.010)	(0.018)
PIT		-0.052***			
		(0.012)			
Social Contributions		0.014			
		(0.010)			
CIT		0.008			
		(0.020)			
Consumption & Property					
Goods & Services			0.047***		0.018
			(0.011)		(0.018)
Trade			0.029		
			(0.018)		
Property			0.048	-0.003	0.019
			(0.032)	(0.028)	(0.031)
LI Dummy	0.016	0.019*	-0.088***	0.028**	0.023
	(0.011)	(0.011)	(0.027)	(0.013)	(0.022)
LMI Dummy	-0.017**	-0.016**	0.034**	-0.017**	-0.012
	(0.008)	(0.008)	(0.013)	(0.008)	(0.018)
UMI Dummy	-0.017**	-0.004	0.016	-0.018**	-0.010
	(0.007)	(0.008)	(0.010)	(0.007)	(0.017)
*LI Dummy*Income*	-0.078**			-0.102***	-0.111***
	(0.033)			(0.037)	(0.042)
*LMI Dummy*Income*	0.048**			0.051***	0.045*
	(0.019)			(0.019)	(0.024)
*UMI Dummy*Income*	0.032**			0.028*	0.026
	(0.015)			(0.015)	(0.022)
*LI Dummy*PIT*		-0.116**			
		(0.051)			
*LMI Dummy*PIT*		0.044**			
		(0.023)			
*UMI Dummy*PIT*		0.020			
		(0.020)			
*LI Dummy*SC*		-0.828**			
		(0.391)			
*LMI Dummy*SC*		0.025			
		(0.043)			
*UMI Dummy*SC*		-0.037*			
		(0.021)			
*LI Dummy*CIT*		-0.080			
		(0.077)			
*LMI Dummy*CIT*		0.073			
		(0.032)			
*UMI Dummy*CIT*		-0.041			
		(0.031)			
*LI Dummy*G&S*			0.134***		0.023
			(0.040)		(0.025)
*LMI Dummy*G&S*			-0.048**		-0.002
			(0.022)		(0.023)
*UMI Dummy*G&S*			-0.036**		-0.011
			(0.017)		(0.021)
*LI Dummy*Trade*			0.111***		
			(0.042)		
*LMI Dummy*Trade*			-0.045*		
			(0.024)		
*UMI Dummy*Trade*			-0.026		
			(0.021)		
*LI Dummy*Property*			-0.418	-0.500	-0.529
			(0.377)	(0.390)	(0.397)
*LMI Dummy*Property*			-0.171	-0.074	-0.126
			(0.140)	(0.140)	(0.143)
*UMI Dummy*Property*			0.056	0.118	0.081
			(0.084)	(0.075)	(0.076)
Omitted Tax Variable	**Consumption & Property**		**Income**	**Consumption**	**Trade**
Observations	2657	2063	2594	2650	2594

*Standard errors in parentheses. *P>0.1; **P>0.05 ;***P>0.01*

The results of the Wald Test for each specification (denoted 'Column 1' – 'Column 5') are provided in Table 4.5. I fail to reject H_0 in column 1, suggesting that the heterogeneous effects of revenue neutral increases in income tax across income groups are indeed statistically different from one another (this result also emerges in columns 4 and 5, where income taxes are included). When income taxes are disaggregated into PIT, social contributions and CIT in column 2, the coefficient estimates on CIT are statistically different from one another at the 1% level, whilst those on PIT and social contributions are statistically different at the 5% level. In column 3, the coefficients on taxes on goods and services are, again statistically different from one another. The same is true for taxes on trade. However, these tests suggest that the effects of property tax on GDP growth rates are not statistically different from one another across subsamples.

Table 4.5 *Results of Wald test*

Column	Column 1	Significance
1	(1) Low income dummy * Income tax share = 0 (2) Lower middle income dummy * Income tax share = 0 (3) Upper middle income dummy * Income tax share = 0 F (3, 2634) = 5.87 Prob > F = 0.0005	***
2	(1) Low income dummy * Corporate income tax share = 0 (2) Lower middle income dummy * Corporate income tax share = 0 (3) Upper middle income dummy * Corporate income tax share = 0 F (3, 2030) = 3.85 Prob > F = 0.0092	***
2	(1) Low income dummy * Personal income tax share = 0 (2) Lower middle income dummy * Personal income tax share = 0 (3) Upper middle income dummy * Personal income tax share = 0 F (3, 2030) = 3.06 Prob > F = 0.0272	**
2	(1) Low income dummy * Social Contributions share = 0 (2) Lower middle income dummy * Social Contributions share = 0 (3) Upper middle income dummy * Social Contributions share = 0 F (3, 2030) = 2.62 Prob > F = 0.0493	**
3	(1) Low income dummy * Goods & Services share = 0 (2) Lower middle income dummy * Goods & Services share = 0 (3) Upper middle income dummy * Goods & Services share = 0 F (3, 2561) = 7.11 Prob > F = 0.0001	***
3	(1) Low income dummy * Trade share = 0 (2) Lower middle income dummy * Trade share = 0 (3) Upper middle income dummy * Trade share = 0 F (3, 2561) = 5.70 Prob > F = 0.0007	*******
3	(1) Low income dummy * Property tax share = 0 (2) Lower middle income dummy * Property tax share = 0 (3) Upper middle income dummy * Property tax share = 0 F (3, 2561) = 1.16 Prob > F = 0.3236	
4	(1) Low income dummy * Income tax share = 0 (2) Lower middle income dummy * Income tax share = 0 (3) Upper middle income dummy * Income tax share = 0 F (3, 2622) = 6.23 Prob > F = 0.0003	*******

4	(1) Low income dummy * Property tax share = 0 (2) Lower middle income dummy * Property tax share = 0 (3) Upper middle income dummy * Property tax share = 0 F (3, 2622) = 1.50 Prob > F = 0.2120	
5	(1) Low income dummy * Income tax share = 0 (2) Lower middle income dummy * Income tax share = 0 (3) Upper middle income dummy * Income tax share = 0 F (3, 2561) = 5.70 Prob > F = 0.0007	***
5	(1) Low income dummy * Tax on Goods & Services share = 0 (2) Lower middle income dummy * Tax on Goods & Services share = 0 (3) Upper middle income dummy * Tax on Goods & Services share = 0 F (3, 2561) = 0.93 Prob > F = 0.4230	
5	(1) Low income dummy * Property tax share = 0 (2) Lower middle income dummy * Property tax share = 0 (3) Upper middle income dummy * Property tax share = 0 F (3, 2561) = 1.32 Prob > F = 0.2676	

*Standard errors in parentheses. *P>0.1; **P>0.05; ***P>0.01*

Thus, the Wald test results are indicative that the main differences in subsamples discussed above are in fact statistically significant, with the exception of property taxes.

4.5.2 Testing validity of Parameter restrictions

The PMG approach used to estimate equation [13] allows for heterogeneous short run effects across countries, but constrains the long run coefficients to be equal. That is, it assumes that the long run relationship between GDP growth and the independent variables is the same across countries. The assumption of long run parameter homogeneity (i.e. that all countries in the sample grow in a similar fashion over time) might be valid for similar groups of countries, such as OECD, or those that form a trading bloc etc. but may not hold across the sample as a whole. It is, however, possible to test the validity of this assumption using the Hausman test to compare the PMG estimates with alternative options. The Mean Group estimator (MG) allows for full parameter heterogeneity; that is, a separate regression is estimated for each group (country) and an average reported. At the other end of the scale, dynamic fixed effects estimation (DFE) constrains all short and long run coefficients to be equal across countries. The estimator employed here, PMG, lies between the two, allowing short run dynamics to vary across countries, whilst constraining the long run coefficients to be equal. Table 4.6 summarises:

Table 4.6 *Parameter restrictions for different estimators*

Estimator	***SR Coefficients***	***LR Coefficients***
Mean Group (**MG**)	Heterogeneous	Heterogeneous
Pooled Mean Group (**PMG**)	Heterogeneous	Homogenous
Dynamic Fixed Effects (**DFE**)	Homogenous	Homogenous

Table 4.7 displays the coefficients of the tax share variables estimated via both PMG and MG (the different specifications here pertain to those in table 4.2) and the resulting Hausman test statistic. In all specifications, the results suggest that the PMG estimator is preferred over the MG. Thus the restriction of parameter homogeneity appears valid, with the PMG procedure producing estimates that are both efficient and consistent.[97] Furthermore, the coefficient estimates for some of tax share variables in the MG regressions seem implausibly high. When tested against the DFE estimator, the Hausman test suggested that the PMG estimator be preferred (for the sake of brevity these results are not shown.) Thus it appears that the assumptions underlying the PMG approach are satisfactory for all specifications.[98]

Table 4.7 *Hausman test: MG v PMG*

	MG	**PMG**
Specification 1		
Income taxes	-0.174	-0.062***
	(0.103)	(0.018)
Hausman Test: Chi2 (5): 1.08; P=0.956		
Specification 2		
PIT	-0.393	-0.068**
	(0.230)	(0.03)
SC	-2.627	-0.090***
	(2.263)	(0.033)
CIT	0.297	0.045
	(0.157)	(0.031)
Hausman Test: Chi2 (7): 2.16; P=0.950		
Specification 3		
GS	0.701	0.069***
	(0.486)	(0.021)
Trade	0.231	0.034
	(0.417)	(0.022)
Property	-12.977	0.058
	(11.745)	(0.076)
Hausman Test: Chi2 (7): 0.054; P=0.999		
Specification 4		
Income	-0.114	-0.047***
	(0.093)	(0.021)
Property	-16.102	-0.015
	(0.015)	(0.077)
Hausman Test: Chi2 (6): 0.76; P=0.993		
Specification 5		
Income	-0.249	-0.034
	(0.422)	(0.022)
GS	0.459	0.036***
	(0.676)	(0.012)
Property	-13.025	0.024
	(11.696)	(0.076)
Hausman Test: Chi2 (7): 3.59; P=0.825		

[97] One caveat that should be noted, however, is that the power of the Hausman Test in this case (i.e. comparing MG and PMG) is relatively low (Pesaran, 1999).

[98] An earlier version of this study (McNabb and LeMay-Boucher, 2014) employd the mean group estimator, however this choice was driven by data availability at the time. The latest version of the GRD contains a substantially more complete series for many countries, thus making estimation via PMG possible.

4.5.3 Alternative time controls

The benchmark results presented above include five year dummies as time controls, in order to capture the effects of the business cycle.[99] Work by Xing (2011), which challenged the results presented in Arnold *et al.* (2011), suggested that results from the PMG estimator may be sensitive to how the time controls are specified. An alternative approach, taken by Acosta-Ormaechea & Yoo (2012) is to include country-specific linear time trends. A replication of table 4.2, using a linear time trend instead of 5-year dummies is included in Appendix G. The majority of results presented above are robust to this alternative time control, with most differences occurring only in the magnitude of coefficient. However, the results presented in tables 4.2 and 4.3 are strongly preferred; it is the standard approach, where growth is the dependent variable, to include some control for the business cycle – a linear trend cannot do this as effectively as the five-year dummy variables.

4.5.4 Addressing Potential Endogeneity concerns

The primary concern with regards endogeneity in this type of model results from the fact that changes in the tax level, or indeed the tax structure, might arise from changes in GDP growth rates. This study does not attempt to ascribe any interpretation on the tax ratio variable, for not only this reason, but also those outlined above in section 4.2. However, it is necessary to (attempt to) rule out the possibility that changes in the tax structure are driven by changes in GDP growth rates. Considering the regression framework here, simultaneity bias might not appear to be a large concern. The dependent variable is the growth rate of log GDP per capita, from *t-1* to *t*; the independent variables are all measured at *t-1*. It is thus unclear how the rate of growth in a future period might drive the share of revenue from a certain tax in the previous year.

A potential source of endogeneity arises from the fact that different taxes' share of total revenue (i.e. the variables of interest here) may react to a change in the level of economic activity in different sectors, which result from some reason other than a change in the tax rate. For example, the share of taxes collected from trade may increase relative to other categories simply as a result of an increase in the volume of trade, regardless of the rate of the taxes levied on imports. In turn, this will also affect GDP growth. The volume of trade openness is included as an additional control in tables 4.8 & 4.9. This is calculated,

[99] These are specified as 1980-84, 1985-89, … ..., 2010-2014.

following Arnold (2008), by obtaining the residuals from a regression of the volume of trade (the sum of the value of imports and exports, expressed as a percentage of GDP) on log population. This therefore represents the part of trade that is not simply explained by country size. For the sake of brevity, only the base specification where consumption and property taxes are excluded and that excluding trade taxes are shown (pertaining to columns 1 and 5 of table 4.2).

All of the aforementioned results hold, with any changes purely in the magnitude of the coefficients. Column 4 of table 4.8 suggests that RN increases in income taxes, offset by reductions in trade taxes are now statistically negatively related to long run GDP growth rates. Turning to table 4.9 column 4, the results now suggest that for high income countries, RN increases in domestic consumption, offset by decreases in trade taxes are harmful for growth. In upper middle income countries, the coefficient on property taxes becomes insignificant (column 8).

Table 4.8 *Full sample, including openness to trade.*

Dependent Variable: Δ(log) GDP per capita.

	1	2	3	4
Physical Capital	0.013	0.009	0.058***	0.078***
	(0.019)	(0.017)	(0.020)	(0.019)
Human Capital	-0.030***	-0.048***	-0.033***	-0.036***
	(0.009)	(0.014)	(0.008)	(0.013)
Population Growth	-0.423***	0.467**	-0.504***	0.777***
	(0.133)	(0.208)	(0.152)	(0.210)
Tax / GDP	0.061**	0.070**	0.124***	0.149***
	(0.029)	(0.027)	(0.032)	(0.028)
Tax Structure Variables				
Income	-0.062***	-0.105***	-0.034	-0.075***
	(0.018)	(0.018)	(0.022)	(0.0230
PIT				
Social Contributions				
CIT				
Consumption & Property				
Goods & Services			0.036***	0.032***
			(0.012)	(0.012)
Trade				
Property			0.024	0.096
			(0.076)	(0.077)
Openness		0.002		-0.001
		(0.001)		(0.002)
Omitted Tax Variable	**Consumption & Property**		**Trade**	
Observations	2657	2657	2290	2274
Groups	100	100	84	83
5-Year Dummies	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Table 4.9 *Results by Income Group, including openness to trade.*

	High Income				Upper- Middle Income			
	1	2	3	4	5	6	7	8
Physical Capital	-0.082**	-0.065*	0.017	0.044	-0.141**	-0.051*	-0.091**	-0.003
	(0.038)	(0.039)	(0.038)	(0.039)	(0.039)	(0.027)	(0.036)	(0.032)
Human Capital	-0.140***	-0.080***	-0.100	-0.038	0.640**	-0.024	0.067***	0.094***
	(0.019)	(0.025)	(0.020)	(0.026)	(0.025)	(0.033)	(0.025)	(0.035)
Population Growth	-0.632***	-0.097	-0.830***	0.301***	1.573***	0.869***	2.049***	1.915***
	(0.190)	(0.357)	(0.212)	(0.373)	(0.265)	(0.319)	(0.220)	(0.337)
Tax / GDP	0.213**	0.174***	0.195***	0.150***	0.142**	-0.129***	0.239***	0.201***
	(0.048)	(0.048)	(0.047)	(0.047)	(0.056)	(0.048)	(0.054)	(0.050)
Tax Structure Variables								
Income taxes	-0.101***	-0.104***	-0.122*	-0.135**	-0.154***	-0.223***	-0.142***	-0.173**
	(0.035)	(0.036)	(0.068)	(0.068)	(0.033)	(0.032)	(0.036)	(0.036)
PIT								
Social Contributions								
CIT								
Consumption & Property								
Goods & Services			-0.099	-0.124**			-0.022	-0.007
			(0.064)	(0.063)			(0.034)	(0.032)
Trade								
Property			0.171	0.128			-0.487***	-0.269
			(0.112)	(0.114)			(0.181)	(0.162)
Openness		-0.007**		-0.005		-0.006		-0.739**
		(0.003)		(0.003)		(0.296)		(0.320)
Omitted Tax Variable	**Consumption & Property**		**Trade**		**Consumption & Property**		**Trade**	
Observations	1202	1202	1121	1121	695	695	663	637
Groups	42	42	38	38	27	27	26	24
5-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Table 4.9 *(Continued)*

	Lower-Middle Income				Low Income			
	9	10	11	12	13	14	15	16
Physical Capital	0.117***	0.168***	-0.016	0.105***	0.152***	0.176***	0.213***	0.158**
	(0.027)	(0.029)	(0.016)	(0.028)	(0.056)	(0.061)	(0.064)	(0.037)
Human Capital	-0.039***	0.020	-0.004	-0.052***	0.044*	-0.115***	0.043	-0.130***
	(0.012)	(0.017)	(0.006)	(0.012)	(0.025)	(0.041)	(0.028)	(0.039)
Population Growth	-1.133***	-0.265	0.421***	-0.365	0.049	0.475	0.421	-3.010***
	(0.295)	(0.428)	(0.101)	(0.253)	(0.303)	(0.579)	(0.346)	(0.447)
Tax / GDP	-0.024	0.050	0.071***	0.119***	-0.241*	-0.156	-0.313**	0.226**
	(0.041)	(0.050)	(0.020)	(0.041)	(0.132)	(0.127)	(0.142)	(0.111)
Tax Structure Variables								
Income	0.001	0.029	0.081***	0.062**	-0.120**	-0.138***	-0.193***	-0.258***
	(0.030)	(0.027)	(0.023)	(0.025)	(0.052)	(0.049)	(0.064)	(0.032)
PIT								
Social Contributions								
CIT								
Consumption & Property								
Goods & Services			0.040***	0.050***			0.011	-0.051***
			(0.011)	(0.010)			(0.026)	(0.020)
Trade								
Property			-1.028***	-0.865***			-0.465*	-0.554***
			(0.274)	(0.282)			(0.247)	(0.222)
Openness		-0.933***		0.001		1.203***		1.309
		(0.244)		(0.002)		(0.345)		(0.231)
Omitted Tax Variable	**Consumption & Property**		**Trade**		**Consumption & Property**		**Trade**	
Observations	359	359	359	359	386	386	383	383
Groups	15	15	15	15	15	15	15	15
5-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

The regression framework employed here includes short run dynamics and 5-year dummies, both of which should aid in accounting for effects of the business cycle. However, a further way in which it is possible to test for the presence of endogeneity (outlined in Acosta-Ormaechea and Yoo (2012), which itself follows an approach outlined in Calderon *et al.* (2011)) is as follows. In order to test if the tax variables considered here are weakly exogenous, the following system of equations is estimated separately for each country, *i,* included in the specifications above.[100]

$$\Delta T_{i,t} = \phi_i \left(g_{i,t-1} - \hat{\alpha}_1 I_{i,t-1} - \hat{\alpha}_2 h_{i,t-1} - \hat{\alpha}_3 n_{i,t-1} - \hat{\alpha}_4 T_{i,t-1} - \sum \hat{\alpha}_j \, TC_{i,t-1} \right) + \varepsilon_{i,t}$$

$$\Delta INC = \phi_i \left(g_{i,t-1} - \hat{\alpha}_1 I_{i,t-1} - \hat{\alpha}_2 h_{i,t-1} - \hat{\alpha}_3 n_{i,t-1} - \hat{\alpha}_4 T_{i,t-1} - \sum \hat{\alpha}_j \, TC_{i,t-1} \right) + \varepsilon_{i,t}$$

$$\Delta GS_{i,t} = \phi_i \left(g_{i,t-1} - \hat{\alpha}_1 I_{i,t-1} - \hat{\alpha}_2 h_{i,t-1} - \hat{\alpha}_3 n_{i,t-1} - \hat{\alpha}_4 T_{i,t-1} - \sum \hat{\alpha}_j \, TC_{i,t-1} \right) + \varepsilon_{i,t} \qquad [15]$$

$$\Delta TRADE_{i,t} = \phi_i \left(g_{i,t-1} - \hat{\alpha}_1 I_{i,t-1} - \hat{\alpha}_2 h_{i,t-1} - \hat{\alpha}_3 n_{i,t-1} - \hat{\alpha}_4 T_{i,t-1} - \sum \hat{\alpha}_j \, TC_{i,t-1} \right) + \varepsilon_{i,t}$$

$$\Delta PROP_{i,t} = \phi_i \left(g_{i,t-1} - \hat{\alpha}_1 I_{i,t-1} - \hat{\alpha}_2 h_{i,t-1} - \hat{\alpha}_3 n_{i,t-1} - \hat{\alpha}_4 T_{i,t-1} - \sum \hat{\alpha}_j \, TC_{i,t-1} \right) + \varepsilon_{i,t}$$

ϕ_i represents the error correction component and each of the terms in parentheses are the long run equilibrium errors resulting from the estimation of equation [13]. This system of equations is estimated by Zellner's (1962) seemingly unrelated regression equations (SURE) method, via the *sureg* command in Stata. For weak exogeneity to hold, it is required that the ϕ_i coefficients are not significantly different from zero. A Wald test is carried out following the SURE regression for each country. If the null hypothesis (that the coefficients on ϕ_i are jointly zero) is rejected at the 5% level, then this suggests that the tax variables under consideration (i.e. the left hand side variables in equation [15]) do in fact react to deviations from the long run relationship (Acosta Ormaechea and Yoo, 2012). As such the weak exogeneity condition is violated in these countries. Depending on the specification tested, between 17 and 24 countries violate the condition of weak exogeneity. Tables 4.10 and 4.11 replicate the results of tables 4.2 and 4.3 respectively, omitting those countries where the tax policy variables cannot be considered weakly exogenous. It was

[100] The number of equations estimated depends on the exact specification of equation [13] being estimated.

not possible to replicate all specifications by country group in table 4.11; due to the smaller *N* dimension, the PMG estimator did not always converge.

The results in table 4.10 are largely in line with those in table 4.2, although some differences are notable. The coefficient on income taxes in column 1 is smaller in magnitude and no longer statistically significant. In column 2, the finding that RN increases in personal income taxes, offset by reductions in consumption or property taxes are statistically negatively associated with long run GDP growth rates remains. However, the coefficient on social contributions is no longer significant and interestingly, there again appears to be some evidence that RN increases in CIT actually have positive effects on long run growth rates.[101] The result in column 5, that RN increases in domestic consumption taxes offset by decreases in trade taxes is good for growth, also holds following the exclusion of the potentially endogenous countries. However, the coefficient estimate is again lower.

Turning to the high income countries in table 4.11, the results suggest that RN shifts toward income taxes are no longer significantly negatively related with long run GDP growth. The results that RN shifts toward property taxes, from either consumption or income taxes, are positively associated with LR growth rates, all remain.

The findings for upper-middle income countries are very similar to those in table 4.3; again, RN shifts toward property taxes appear to be negatively related with long run GDP growth rates (column 7), although the coefficient is somewhat smaller. Columns 9-12 show that the results for lower-middle income countries also remain robust; the key finding, that RN increases in domestic consumption taxes offset by decreases in trade taxes are positively associated with long run GDP growth rates, still holds, though the coefficient estimate is now over twice that reported in table 4.3. There are no notable differences observed between table 3 and table 10 for low income countries, aside from small variations in some coefficient estimates and the significance level of some of the tax variables in column 14.

[101] However, this result does not hold following further checks where both (i) the potentially endogenous countries were excluded and (ii) the openness variable was included. Thus it is difficult to draw any concrete conclusions from this particular result.

Table 4.10 *Replication of Table 4.2, excluding countries identified as potentially endogenous*

Dependent Variable: Δ(log) GDP per capita.

	1	2	3	4	5
Physical Capital	0.030	0.032	-0.012	0.050**	0.068***
	(0.020)	(0.025)	(0.017)	(0.021)	(0.022)
Human Capital	-0.041***	-0.025*	0.007	-0.031***	-0.041***
	(0.012)	(0.013)	(0.007)	(0.011)	(0.012)
Population Growth	-0.415**	-0.163	0.425***	-0.628***	-0.602***
	(0.133)	(0.193)	(0.108)	(0.165)	(0.180)
Tax / GDP	0.025	0.017	0.098***	0.081***	0.070**
	(0.031)	(0.036)	(0.022)	(0.031)	(0.034)
Tax Structure Variables					
Income	-0.020			-0.040*	-0.031
	(0.021)			(0.023)	(0.026)
PIT		-0.086**			
		(0.037)			
Social Contributions		-0.032			
		(0.042)			
CIT		0.146***			
		(0.039)			
Consumption & Property					
Goods & Services			0.012		0.026***
			(0.024)		(0.017)
Trade			-0.039		
			(0.024)		
Property			0.189*	0.114	0.164
			(0.098)	(0.092)	(0.099)
Omitted Tax Variable	**Consumption & Property**		**Income**	**Consumption**	**Trade**
Observations	2001	1273	1719	1812	1745
Groups	76	48	63	66	64
5-Year Dummies	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Table 4.11 *Replication of Table 4.3, excluding countries identified as potentially endogenous*

Dependent Variable: Δ(log) GDP per capita.	**High Income**				**Upper-Middle Income**			
	1	2	3	4	5	6	7	8
Physical Capital	-0.060	0.039	-0.025	0.027	-0.132***	-0.003	-0.099**	0.003
	(0.045)	(0.043)	(0.043)	(0.044)	(0.042)	(0.042)	(0.049)	(0.042)
Human Capital	-0.118***	-0.085***	-0.121***	-0.089	0.057*	0.110**	0.074	0.110**
	(0.025)	(0.022)	(0.022)	(0.022)	(0.030)	(0.055)	(0.046)	(0.055)
Population Growth	-0.617**	-0.937***	-0.897***	-0.923***	1.801***	2.248***	2.069***	2.248***
	(0.264)	(0.240)	(0.237)	(0.244)	(0.269)	(0.205)	(0.275)	(0.205)
Tax / GDP	0.142**	0.120**	0.184***	0.109**	0.147**	0.308***	0.191***	0.308***
	(0.057)	(0.054)	(0.052)	(0.054)	(0.057)	(0.065)	(0.073)	(0.065)
Tax Structure Variables								
Income taxes	-0.307		0.016	-0.028	-0.131***		-0.227***	-0.261***
	(0.040)		(0.040)	(0.078)	(0.036)		(0.061)	(0.055)
PIT								
Social Contributions								
CIT								
Consumption & Property								
Goods & Services		0.009		0.022		0.173***		-0.088*
		(0.039)		(0.073)		(0.061)		(0.047)
Trade		0.059				0.261***		
		(0.076)				(0.055)		
Property		0.393***	0.471***	0.442***		0.549**	-0.149***	0.288***
		(0.125)	(0.124)	(0.152)		(0.268)	(0.227)	(0.264)
Omitted Tax Variable	Consumption & Property	Income	Consumption	Trade	Consumption & Property	Income	Consumption	Trade
Observations	902	850	921	834	511	288	289	288
Groups	32	29	31	29	20	11	11	11
5-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

Table 4.11 *(continued)*

Dependent Variable: Δ(log) GDP per capita.	Lower-Middle Income				Low Income			
	9	10	11	12	13	14	15	16
Physical Capital	0.118***	-0.011	0.117***	-0.009	0.144**	0.282***	0.156**	0.266***
	(0.027)	(0.022)	(0.026)	(0.022)	(0.062)	(0.074)	(0.066)	(0.068)
Human Capital	-0.045***	-0.011	-0.025	-0.005	0.091***	0.057	0.070**	0.026
	(0.015)	(0.008)	(0.013)	(0.008)	(0.034)	(0.039)	(0.034)	(0.028)
Population Growth	-1.511***	0.379***	-1.642***	0.416***	-0.141	0.287	0.356	0.706*
	(0.355)	(0.103)	(0.351)	(0.118)	(0.477)	(0.507)	(0.469)	(0.419)
Tax / GDP	-0.024	0.083***	0.022	0.106***	-0.259	-0.241	-0.229	-0.225
	(0.044)	(0.026)	(0.042)	(0.028)	(0.165)	(0.181)	(0.167)	(0.152)
Tax Structure Variables								
Income taxes	0.016		-0.018	0.104***	-0.170**		-0.126*	-0.205***
	(0.036)		(0.030)	(0.029)	(0.068)		(0.071)	(0.067)
PIT								
Social Contributions								
CIT								
Consumption and Property								
Goods & Services		-0.070**		0.084***		0.156*		0.011
		(0.026)		(0.024)		(0.084)		(0.026)
Trade		-0.131***				0.125		
		(0.027)				(0.080)		
Property		-0.722***	-0.787***	-0.579**		0.623*	0.281	-0.485**
		(0.279)	(0.259)	(0.289)		(0.312)	(0.304)	(0.247)
Omitted Tax Variable	Consumption & Property	Income	Consumption	Trade	Consumption & Property	Income	Consumption & Property	Trade
Observations	304	268	291	290	275	304	307	333
Groups	13	11	12	12	11	12	12	13
5-Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

4.5.5 Accounting for Cross Sectional Dependence

A further source of bias that might arise in a macro panel such as that used here is via residual cross-section dependence (CSD). This occurs when unobserved common shocks affect all countries or a subset of countries in the dataset. In our context, such shocks might take the form of (e.g.) commodity price fluctuations or tax agreements where a number of countries agree to reduce tariffs on each others' imports. As proposed by Pesaran (2006), a simple way to account for the existence of CSD is to augment the equation being estimated with cross-sectional averages of the dependent and independent variables, i.e.

$$\frac{1}{N}\sum_{i=1}^{N}\Delta g_{it} \qquad \& \qquad \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{X}_{it}$$

respectively, where $\boldsymbol{X}_{it}$ is the vector of all explanatory variables. Given that the PMG estimator uses maximum likelihood, augmenting the estimated equation with *k* cross-sectional averages can lead to difficulties in the estimation procedure (the estimator might fail to converge or the likelihood function become non-concave). Thus it was only possible to fully replicate the results of table 4.2 and indeed for many specifications the sample size is reduced as the estimator only converges when those countries with a sufficiently long *t* dimension are included in the analysis. These results are shown in table 4.12.

The result in column 1 remains unchanged and the coefficient estimate on the income tax share is almost identical to that in table 4.2. It was not possible to repeat the estimations with income tax disaggregated into PIT, Social Contributions and CIT. In column 2, the results show that RN shifts toward property taxes, away from income taxes, are most growth-friendly, followed by RN shifts toward trade taxes. The result in table 4.2, that RN shifts away from income toward taxes on goods and services led to higher growth rates is no longer statistically significant. Results in column 3 suggest that RN shifts toward income, and away from consumption, taxes have negative effects on GDP growth rates; notably however the coefficient estimate of -0.082 is somewhat larger than that reported in table 4.2 (-0.047). Finally, in column 4, RN shifts away from trade taxes toward domestic consumption taxes are no longer significantly positively related to GDP. However, RN increase in property taxes are.

Thus, after attempting to control for cross-sectional averages, there are a couple of differences with the benchmark results. However, a potential issue with this set of results

is that the sample size is reduced for each estimation; it is therefore difficult to say whether the difference in results is due to the inclusion of the cross-sectional averages, or due to the smaller sample. It is hoped that in future work these results will become more clear via use of the dynamic common correlated effects estimator (Ditzen, 2016), which is to be released in the near future. This should avoid the issues with convergence experienced with *xtpmg* as it uses OLS in order to estimate pooled mean-group regressions, allowing for the effects of CSD to be investigated in the full panel of countries.

Table 4.12. *Replication of Table 4.2, after including cross-sectional averages of all variables*

Dependent Variable: Δ(log) GDP per capita.

	1	2	3	4
Physical Capital	-0.004	0.046***	0.026	0.037*
	(0.017)	(0.022)	(0.018)	(0.023)
Human Capital	-0.005	-0.088***	-0.092***	-0.077***
	(0.018)	(0.024)	(0.022)	(0.025)
Population Growth	-0.297**	-0.545***	0.032	-0.707***
	(0.140)	(0.186)	(0.168)	(0.184)
Tax / GDP	0.138***	0.066*	0.133***	0.075**
	(0.030)	(0.039)	(0.034)	(0.038)
Tax Structure Variables				
Income	-0.059***		-0.082***	-0.017
	(0.017)		(0.018)	(0.027)
PIT				
Social Contributions				
CIT				
Consumption & Property				
Goods & Services		0.011		-0.020
		(0.024)		(0.024)
Trade		0.056**		
		(0.028)		
Property		0.247***	0.006	0.190***
		(0.072)	(0.067)	(0.071)
Omitted Tax Variable	**Income**		**Consumption**	**Trade**
Observations	2388	1777	2040	1777
Groups	85	58	70	58
5-Year Dummies	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

4.5.6 Excluding resource-rich countries.

As highlighted in figure 4.4, there are a number of resource-rich (specifically, oil-producing) countries where the tax ratio is very low, but per capita income is high. Thus in order to ensure that the results reported previously are not biased by the inclusion of resource rich countries where growth (and also tax revenues) can fluctuate wildly based on, for example, commodity prices, the specifications above are re-run excluding resource-rich countries. These were identified by eyeballing the data in the GRD for (i) countries reporting nontax revenues above 10% of GDP and (ii) countries reporting a high level of

resource tax revenues. Those dropped from the sample above are Algeria, Bahrain, Botswana, Egypt, Gabon, Morocco, Sudan, Trinidad and Tobago and Venezuela. Table 4.13 reports the results for the full sample with these countries excluded.

Table 4.13 *Replication of Table 4.2, excluding resource-rich countries*

Dependent Variable: Δ(log) GDP per capita.

	(1)	(2)	(3)	(4)	(5)
Physical Capital	0.022	0.030	-0.030*	0.039*	0.046***
	(0.020)	(0.022)	(0.016)	(0.021)	(0.022)
Human Capital	-0.027***	-0.047***	0.011**	-0.028***	-0.031***
	(0.009)	(0.011)	(0.005)	(0.009)	(0.009)
Population Growth	-0.493***	-0.538***	0.371***	-0.600***	-0.441***
	(0.137)	(0.160)	(0.106)	(0.148)	(0.158)
Tax / GDP	0.046	(0.019)**	0.100***	0.095***	0.135***
	(0.030)	(0.008)	(0.020)	(0.031)	(0.033)
Tax Structure Variables					-0.070***
					(0.024)
Income taxes	-0.097***			-0.074***	
	(0.021)			(0.023)	
PIT		-0.052*			
		(0.029)			
Social Contributions		-0.092***			
		(0.033)			
CIT		0.048			
		(0.030)			
Consumption & Property					
Goods & Services			0.066***		0.042***
			(0.022)		(0.012)
Trade			0.015		
			(0.021)		
Property			0.039	-0.019	0.003
			(0.083)	(0.080)	(0.080)
Omitted Tax Variable	**Consumption & Property**		**Income**	**Consumption**	**Trade**
Observations	2413	1804	2134	2134	2134
Groups	91	68	78	78	78
5-Year Dummies	Yes	Yes	Yes	Yes	Yes

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*

All of the previously reported results hold, with similar coefficient estimates. Notably, however, the result in column 1 suggests a somewhat stronger effect of RN increases in income tax, offset by consumption and property taxes; a percentage point increase in income taxes leads to a 0.097 percent reduction in the long run growth rate.

4.5.7 Further limitations

There are a number of further limitations to a study of this nature. Firstly, it is inherently difficult to account for the fact that changes in tax policy are often announced some time in advance of implementation. Individuals and firms may therefore adjust their behaviour (which in turn influences the share of tax from any one component) before the tax rate itself has actually changed. A further challenge lies in the fact that it is difficult to account for (changes in) the efficiency of tax collection, which has clear implications for GDP growth:

even a well-designed tax system might be undermined by poor administration. However, the human capital variable might go some way to capturing citizens' ability to understand and comply with tax laws, and the inclusion of the tax level (as a share of GDP), to an extent, serves as a control for the ability of the government to administer the tax system.[102] Parallel to difficulty in accounting for the *ability* of a country to collect different taxes, it is also difficult to control for the *cost* of collecting different types of taxes. The above analysis, and related studies in the literature, assume that the costs of collecting different taxes are equal. This is, of course, unlikely to be true in practice. Similarly, studies of this nature assume that tax *design* does not matter for growth but again, it obviously does. IMF (2013) makes this clear, noting the difference between a corporate tax on rents and one on total returns; the former would not affect the marginal incentive to invest whilst the latter most certainly would. Whilst the assumption of revenue-neutrality makes possible this kind of empirical analysis, there may be limits to the insights presented; many tax authorities in developing countries may altar their tax mix in the hope that the tax ratio itself increases. Furthermore, the magnitudes of the coefficients are quite small, suggesting limited direct effects of any RN changes in the tax mix on long run GDP growth rates. Finally, the distinction between domestic taxes on goods and services and trade taxes is often somewhat blurry, especially in developing countries; whilst one country collecting VAT at the border might classify this as a trade tax, the next may count it as a tax on goods and services. Thus, the results that distinguish between domestic consumption and trade taxes should be interpreted with this in mind. It is hoped that future development of the GRD will seek to improve in this respect.

4.6 Discussion and Conclusion

It is clear from the results presented here that different revenue neutral changes in the tax mix seem to matter to a different extent in different countries. In short, this highlights the pitfalls of promoting a *one size fits all* policy recommendation. This finding is key; Prichard (2016) notes the message that increases in personal and corporate income taxes are bad for growth has become policy orthodoxy at the IMF. One need not look far to see evidence of this – IMF (2011) and IMF (2015) both make reference to the thinking that indirect taxes are more growth friendly than income taxes. But this thinking is based on studies that have been carried out only in high income, or OECD, countries.

[102] That is, assuming that larger governments (as proxied by larger tax/GDP figures) are more able to collect taxes. Of course this says nothing about the *efficiency* of tax collection, but it is plausible that where more people are employed in the public sector, tax collection may be more sophisticated.

A knowledge of those taxes that are most growth friendly at different levels of income can aid policymakers in a way that results from a sample containing only high income countries cannot. Not only has GDP growth over the last 30 years been, on average, lower in high - income countries than in developing countries, but the tax structure has remained remarkably stable. Thus any growth effects of a change in tax structure are likely to be very small in magnitude. However, in low and middle income countries, the past 30 years have seen dramatic changes in the tax mix, with large scale shifts away from a reliance on trade taxes, toward taxes on domestic goods and services and to a lesser extent, income taxes. For many countries, this is likely to continue in the future. Thus the potential to affect long run growth rates with tax policy is much greater in such countries. The results presented here provide several new insights:

Firstly, revenue - neutral reductions in trade taxes, offset by increases in domestic consumption taxes appear to be growth-friendly. Importantly, however, this result seems to be driven by the group of countries classed as *lower-middle income*. No such effect is found for *low-income* countries. Indeed, after controlling for the degree of openness to trade, the results suggest a significant *negative* effect of RN shifts toward taxes on goods and services. Thus one might tentatively conclude that trade liberalisation in the very poorest countries does not lead to higher economic growth, but at later levels of development, it can make a positive contribution to growth rates. When viewed alongside the results of Baunsgaard and Keen (2010), who found revenue recovery to be extremely poor in low income countries that had removed trade barriers, the difficulties of promoting trade liberalisation in the very poorest of countries is further underlined.

Secondly, the results presented here confirm that revenue – neutral shifts away from consumption and property taxes, toward income taxes, are harmful for GDP growth rates, as previously found by, for example, Arnold *et al.* (2011) or Acosta-Ormaechea and Yoo (2012). However again, the magnitude of the effect differs at different levels of development, with the strongest negative effects seen in *upper – middle income* countries and no significant effects seen in *lower-middle income* countries. Personal income taxes and social contributions appear most harmful for long run GDP growth rates, but no evidence is found that increases in corporate income taxes are harmful for growth.

Thirdly, the thinking that revenue neutral increases in property taxes are good for economic growth, as previously suggested in other studies, is called into question. The results suggest

that whilst this may be the case for high-income countries, revenue neutral increases in property taxes in low or middle income countries might have limited or indeed detrimental effects on long run GDP growth rates.

As a result of the much improved data availability in the GRD, this study has been able to extend the existing research (especially in regard to developing countries) and potentially offer insight that is more relevant to developing countries. This is done in a more transparent manner – all of the data used here is publically available, unlike other recent work. If advice on the direction of fiscal policy or the structure of the tax mix is to be provided to developing countries, then at the very least this should be based on evidence of the experience of other developing countries. Given that domestic tax collection is now, on average, the largest form of revenue for developing countries, this work presents an important first step toward providing evidence of how changes in the tax mix might have the potential to affect GDP growth rates.

5. Conclusion

This thesis has considered three facets of economic development, namely (i) the sustainability of informal savings, (ii) primary school attendance and (iii) the role of tax structure in explaining economic growth. It has, for the first two studies, considered data from Benin, a country that has not been studied in great detail in the related literatures. The thesis has also sought to use new, or better, data and methods where appropriate, in order to test existing beliefs and theories.

The first study considered how ROSCAs can be designed in order to minimise the payoff to potential defaulters and thus help to sustain the groups. In a context where formal savings options are limited, often individuals must turn to informal arrangements such as a ROSCA in order to finance the purchase of durable goods, stock for a business, or even to just commit to a savings pattern. However, if the incentives for individuals to cheat on the arrangement outweigh those to remain true, then they are inherently risky and perhaps no more desirable than saving in autarky. Using a unique panel dataset collected in 2004 and 2006 in Cotonou, Benin, the empirical analysis examined the likelihood that a ROSCA experienced a case of default between the two surveys. The group-level estimations present a first empirical test of the theories outlined in studies such as Besley *et al.* (1993) and Anderson *et al.* (2009). The main results highlight that there are certain features of ROSCAs that are associated with a lower likelihood of default. Specifically, those groups run by a president, that met less frequently, that had written rules or were started amongst family members were all significantly less likely to have experienced a case of default in the years between the surveys. Furthermore, the importance of these features diminished as groups successfully completed more cycles. These results are interesting as we look to the future of savings in developing countries; a knowledge of the ROSCA structure that will bring about the lowest likelihood of default is useful to policymakers, microfinance institutions or even formal banks. For example, El-Gamal *et al.* (2014) found that MFIs taking the form of a *bank-insured ROSCA* had higher repayment rates than those following a traditional *Grameen* style. Furthermore, Handa and Kirton (1999) note that the Worker's Bank of Jamaica offers a 'partner account', the structure of which resembles a ROSCA. Thus in the future, more MFIs, NGOs, or even formal banks wishing to connect with new customers might initially offer products that resemble a ROSCA, with which many individuals will be familiar. If this is the case, then knowledge of the ROSCA design that sees higher repayment rates may be of key importance. However, a note of caution might

be urged when interpreting the results; a lack of information on (i) the exact circumstances surrounding each enforcement problem and (ii) each and every member present (usually only one or two representative members were interviewed) means that it is not possible to confidently interpret the results as causal. However, the points raised in section 4.2 give reason to be reasonably confident that the structure of the group is not related to the motivations of members. Future research might wish to take account of these issues.

The analysis in chapter 3 considered the determinants of primary school enrolment in Benin for the 2005-06 school year. In particular, regional and gender disparities were studied in detail. The importance of education in development is well recognised; achieving universal primary education was one of the MDGs and ensuring that both boys and girls complete equitable primary education is embodied in the SDGs. In this respect, Benin provides a stimulating case study. Having recovered from some of the worst primary school enrolment rates in the world in 1990, the country is still faced with considerable regional differences in attendance rates. Furthermore, girls remain significantly less likely to attend school than boys. Chapter 3 highlights that the official statistics on school enrolment in Benin from INSAE or UIS might well suffer from over-reporting; this finding backs up what has previously been reported in other contexts (e.g. Sandefur and Glassman, 2015). Using data from the 2006 Demographic and Health Surveys along with school supply statistics from INSAE, the empirical analysis of chapter 3 firstly seeks to highlight those factors that are associated with the likelihood that a child attends primary school. Boys, birth children, those from richer households and those from Christian households were all significantly more likely to be attending school. An in-depth analysis of the role of the opportunity cost of attending reveals that as distance to school (which represents time spent travelling) increased, this reduced the likelihood that boys who worked in the field would attend to a greater extent than it did for those boys not working. This provides evidence that in a context where households might not be able to fully measure the costs and benefits of sending their children to school, they do in fact consider the opportunity cost. A three-level model is also estimated, which takes account of higher level clustering in the data. This showed that the supply of schools, or distance to school, explained only some of the inter-communal differences in attendance rates, suggesting that regional differences in labour markets, culture or tradition might well provide more insight into the disparities in attendance rates observed. Moreover, the model showed that, after controlling for the observable factors, only 4.9% of the remaining variation in attendance rates was at the commune level, whilst some 30% was at the household level. As such, the largest

improvements in attendance rates might be realised by focusing on factors at the household level, such as raising income or tackling attitudes toward education of daughters. The random-slopes model helped to pinpoint those communes where school attendance rates were below average, but the impact of household wealth on attendance was above average. Were policymakers in Benin interested in reducing regional disparities in primary school attendance, then interventions that helped to raise household income (or lower the cost of schooling) in these communes might be most effective.

The third study examined the relationship between tax structure and economic growth in a panel of low-, middle- and high – income countries. Recent work in this area (e.g. Arnold *et al.,* 2011; Acosta - Ormaechea and Yoo, 2012) has found support for the theory that income taxes are more harmful to GDP growth than consumption, or property taxes. In turn, this evidence has been used to inform the knowledge base which serves to shape policy recommendations at, for example, the IMF. Crucially, the aforementioned studies focus their analysis on OECD or high income countries, whilst those countries most likely to reform their tax structure in the coming years are more likely to be low or middle income countries. Chapter 3 extends work in this field by considering the effects of revenue neutral changes in tax structure on a panel of 100 high, middle and low income countries. The analysis is made possible due to the increased availability of data on tax revenues in the ICTD – UNU-WIDER Government Revenue Dataset. The key finding is that previous results are not generalizable to countries at all income levels. Results suggest that revenue-neutral increases in income taxes, offset by reductions in consumption or property taxes are indeed harmful for economic growth rates, but not across all income levels. This result appears to hold up after eliminating those countries where the tax policy variables were identified as potentially endogenous, accounting for residual cross-sectional dependence and controlling for the degree of openness to trade. The previously reported result (e.g. Arnold *et al.*, 2011) that property taxes are most growth friendly would appear to be true for high income countries, but mixed evidence is reported elsewhere. There is also no evidence in the data used here that corporate income taxes are harmful for growth rates, as theory or some previous empirical work would suggest. Furthermore, for some countries it would appear that the ongoing replacement of trade tariffs with domestic consumption taxes has not has positive effects on GDP growth. Overall, these results highlight the importance of future work in this area paying attention to the heterogeneous effects of tax structure on growth at different stages of development. Whilst these results have called into question, and in some cases overturned, those of previous authors, future research in the

tax-growth domain might seek to pay further attention to the questions of endogeneity and cross sectional dependence. Additionally, future work may seek to relax the revenue-neutrality constraint, which is perhaps unrealistic in reality; some taxes are, naturally, better revenue-raisers than others. Thus an interesting next step may be an attempt to disentangle whether increases in certain types of taxation raise more revenue (which may fuel faster economic growth rates), have direct impacts on economic growth, or a combination of the two.

References

Acosta-Ormaechea, S. & Yoo, J., (2012) 'Tax Composition and Growth: A Broad Cross Country Perspective', IMF Working Paper WP/12/257, International Monetary Fund.

Ahlin, C., (2015) 'The role of group size in group lending', *Journal of Development Economics*, Vol. 115, pp.140-155.

Allen, C., (1989) 'Benin'
in
Allen, C., Radu, M. S. & Somerville, K., (1989) 'Benin, the Congo, Burkina Faso. Politics, Economics and Society', Thomson Learning.

Ambec, S. & Treich, N., (2007) 'Roscas as financial agreements to cope with self-control problems', *Journal of Development Economics*, Vol. 82, pp. 120–137.

Anderson, S. & Baland, J-M., (2002) 'The economics of roscas and intra-household resource allocation', *Quarterly Journal of Economics*, Vol. 117, pp. 963–995.

Anderson, S., Baland, J-M. & Moene, K.O., (2009) 'Enforcement in informal savings groups', *Journal of Development Economics*, Vol. 90, pp. 14-23.

Armendariz, B. & Murdoch, J., (2005) 'The Economics of Microfinance', MIT Press.

Arnold, J., (2008) 'Do Tax Structures affect Aggregate Economic Growth? Empirical evidence from a panel of OECD countries', OECD Working Paper ECO/WKP(2008)51.

Arnold, J., Brys, B., Heady, C., Johansson, Å., Schwellnus, C. & Vartia, L., (2011) 'Tax Policy for Economic Recovery and Growth', *The Economic Journal*, 121(550), pp. F59-F80.

Barro, R., (1990) 'Government Spending in a Simple Model of Endogenous Growth', *Journal of Political Economy*, 98(5), pp. S103-S125.

Barro, R. & Lee, J., (2013) 'A new data set of educational attainment in the world, 1950 - 2010', *Journal of Development Economics*, 104, pp. 184-198.

Barro, R. & Sala-i-Martin, X., (1995) 'Economic growth', New York: McGraw-Hill.

Bartus, T., (2005) 'Estimation of marginal effects using margeff', *Stata journal*, *5*(3), pp. 309-329.

Basu, K., (2011) 'Hyperbolic discounting and the sustainability of rotational savings arrangements', *American Economic Journal: Microeconomics*, Vol. 3(4), pp. 143-171.

Baunsgaard, T. & Keen, M., (2010) 'Tax revenue and (or?) trade liberalization', *Journal of Public Economics*, 94(9), pp. 563-577.

Becker, G.S., (1975) 'Human Capital, A Theoretical and Empirical Analysis, with Special Reference to Education', 2nd Ed., *National Bureau of Economic Research,* Columbia University Press, New York.

Besley, T., Coate, S. & Loury, G., (1993) 'The economics of rotating savings and credit associations', *American Economic Review*, Vol. 83, pp. 792–810.

Biggart, N. W., (2001) 'Banking on each other: the situational logic of rotating savings and credit associations', *Advances in Qualitative Organization Research*, Vol. 3(1), pp. 129-52.

Blackburne, E. & Frank, M., (2007) 'Estimation of Nonstationary Heterogeneous Panels', *The Stata Journal*, 7(2), pp. 197-208.

Bonan, J., LeMay-Boucher, P. & McNabb, K., (2013) 'Are Hyperbolic Agents Sophisticated? An Investigation into Rosca and Funeral Group Membership in Benin', mimeo, Heriot-Watt University.

Bonnett, A.W., (1981) 'Institutional adaptation of West Indian immigrants to America: An analysis of rotating credit associations', Washington DC: University Press of America.

Buchmann, C., (2000) 'Family Structure, Parental Perceptions, and Child Labor in Kenya: What Factors Determine Who Is Enrolled in School?', *Social Forces*, 78(4), pp.1349-1379.

Calerdón, C., Moral-Benito, E. & Servén, L., (2011) 'Is Infrastructure Capital Productive? A dynamic heterogeneous approach', *Policy Research Working Paper Series 5682*, The World Bank.

Callier, P., (1990) 'Informal Finance: The rotating savings and credit association – An interpretation', *Kyklos*, Vol. 43(2), pp. 273-276.

Chernichovsky, D., (1985) 'Socioeconomic and Demographic Aspects of School Enrollment and Attendance in Rural Botswana', *Economic Development and Cultural Change*, 33(2), pp. 319 – 332.

Chiteji, N.S., (2002) 'Promises Kept: Enforcement and the role of rotating savings and credit associations in an economy', *Journal of International Development*, Vol. 14, pp. 393-411.

Clemens, M., (2015) 'The meaning of failed replications: A review and proposal', *Journal of Economic Surveys*, doi:10.1111/joes.12139.

Colclough, C., Rose, P. & Tembon, M., (2000) 'Gender inequalities in primary schooling: The roles of poverty and adverse cultural practice', *International Journal of Educational Development*, 20(1), pp. 5-27.

Coleman, J., (1988) 'Social Capital in the Creation of Human Capital', *American Journal of Sociology*, Supplement 94, pp. S95-S120.

Csapo, M., (1981) 'Religious, Social and Economic Factors Hindering the Education of Girls in Northern Nigeria', *Comparative Education*, 17(3), pp. 311-319.

Dagnelie, O. & LeMay-Boucher, P., (2012) 'Rosca participation in Benin: A commitment issue', *Oxford Bulletin of Economics and Statistics*, Vol. 74(2), pp. 235-252.

Deininger, K., (2003) 'Does cost of schooling affect enrolment by the poor? Universal primary education in Uganda', *Economics of Education Review*, 22, pp. 291-305.

Delprato, M. & Sabates, R., (2015) 'A multilevel analysis of late entry in Nigeria', *International Journal of Research & Method in Education*, 38(1), pp. 83-99.

Devarajan, S., (2013) 'Africa's Statistical Tragedy', *Review of Income and Wealth*, 59(S1), pp. S9-S15.

DHS, (2004) 'The DHS Wealth Index', DHS Comparative Reports, 6,
Available at: <http://www.dhsprogram.com/pubs/pdf/CR6/CR6.pdf> *Accessed 1/3/2015.*

Ditzen, J., (2016) 'Estimating dynamic common correlated effects in Stata', mimeo, Heriot-Watt University.

Djankov, S., La Porta, R., Lopez-de-Silanes, F. & Shleifer, A., (2002) 'The regulation of entry', *The Quarterly Journal of Economics*, 117(1), pp. 1-37.

Dupas, P., Green, S., Keats, A. & Robinson, J., (2012) 'Challenges in Banking the Rural Poor: Evidence from Kenya's Western Province', NBER WORKING PAPER No. 17851.

Easterly, W. & Rebelo, S., (1993) 'Fiscal policy and economic growth', *Journal of Monetary Economics*, 32(3), pp. 417-458.

El-Gamal, M., El-Komi, M., Karlan, D. & Osman, A., (2014) 'Bank-Insured RoSCA for microfinance: Experimental evidence in poor Egyptian villages', *Journal of Economic Behavior and Organization*, 103, pp. S56-S73.

Eloundou-Enyegye, P. M. & Calve`s, A. E., (2006) 'Till marriage do us part: Education and remittances from married women in Africa', *Comparative Education Review*, 50(1), pp. 1–20.

Feigenberg, B., Field, E. & Pande, R., (2013) 'The Economic Returns to Social Interaction: Experimental Evidence from Microfinance', *Review of Economic Studies*, Vol. 80, pp. 1459-1483.

fhi360, (2013) 'Out of School Children. Data Challenges in Measuring Access to Education', *Education Policy and Data Center,* fhi360.
Available at:
<http://www.fhi360.org/sites/default/files/media/documents/Out_of_school_children_data_challenges.pdf> *Accessed 25/6/2015.*

Field, E. & Pande, R., (2008) 'Repayment Frequency and Default in Microfinance: Evidence from India', *Journal of the European Economic Association*, 6(2-3), pp. 501-509.

Ghatak, M., (1999) 'Group Lending, local information and peer selection', *Journal of Development Economics*, Vol. 60, pp. 27-50.

Glick, P., (2008) 'What Policies Will Reduce Gender Schooling Gaps in Developing Countries: Evidence and Interpretation', *World Development*, 36(9), pp. 1623-1646.

Glick, P. & Sahn, D., (2006) 'The demand for primary schooling in Madagascar: Price, Quality, and the choice between public and private providers', *Journal of Development Economics*, 79, pp. 118-145.

Gracia, M., (2000) 'Apport de la Banque Ambulante aux Systèmes Financiers Décentralises et aux Opérateurs Economiques', Paper presented at the International Symposium on Microfinance and the Promotion of Micro- and Small Enterprises, 12–14 June 2000, Praia, Cape Verde.

Gugerty, M. K., (2007) 'You can't save alone: Commitment in Rotating Savings and Credit Associations in Kenya', *Economic Development and Cultural Change*, Vol. 55(2), pp. 251-282.

Gumus, S., (2014) 'The effects of community factors on school participation in Turkey: A multilevel analysis', *International Review of Education*, 60(1), pp. 79-98.

Handa, S. & Kirton, C., (1999) 'The economics of rotating savings and credit associations evidence from the Jamaican "Partner"', *Journal of Development Economics*, Vol. 60, pp. 173–194.

Helms, B., Deshpande, R., Pickens, M. & Sado, N., (2005) 'Benin: country-level savings assessment.*' CGAP Savings Initiative*. Washington, DC: World Bank. <http://documents.worldbank.org/curated/en/2005/08/6726005/benin-country-level-savings-assessment> *Accessed 27/7/2016.*

Huisman, J. & Smits, J., (2009) 'Effects of household-and district-level factors on primary school enrollment in 30 developing countries', *World Development*, *37*(1), pp. 179-193.

IMF, (2011) 'Revenue Mobilization in Developing Countries', *International Monetary Fund,* Washington DC, USA <https://www.imf.org/external/np/pp/eng/2011/030811.pdf> *Accessed 10/02/2016.*

IMF, (2013) 'Taxing Times', Fiscal Monitor, *International Monetary Fund,* October 2013, Washington DC, USA <https://www.imf.org/external/pubs/ft/fm/2013/02/pdf/fm1302.pdf> *Accessed 10/02/2016.*

IMF, (2015) 'Fiscal Policy and Long Term Growth', IMF Policy Paper, *International Monetary Fund*, Washington DC, USA. <http://www.imf.org/external/np/pp/eng/2015/042015.pdf> *Accessed 10/02/2016.*

Jerven, M., (2013) 'Poor numbers: How we are misled by African development statistics and what to do about it', Ithaca: Cornell University Press.

Kandel, E. & Lazear, E. P., (1992) 'Peer pressure and partnerships', *Journal of Political Economy*, Vol. 100(4), pp. 801-817.

Karlan, D., (2007) 'Social Connections and Group Banking', *Economic Journal,* Vol. 117(517), pp. F52-F84.

Kazeem, A., Jensen, L. & Stokes, C. S., (2010) 'School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence, and Socioeconomic Status', *Comparative Education Review*, 54(2), pp. 295-319.

Kedir, A.M. & Ibrahim, J., (2011) 'ROSCAs in Urban Ethiopia: Are the Characteristics of the Institutions More Important than those of Members?', *Journal of Development Studies,* 47(7), pp. 998-1016.

Kneller, R., Bleaney, M. & Gemmell, N., (1999) 'Fiscal policy and growth: evidence from OECD countries', *Journal of Public Economics*, 74(2), pp. 171-190.

Lavy, V., (1996) 'School supply constraints and children's educational outcomes in rural Ghana', *Journal of Development Economics*, 51, pp. 291-314.

Leckie, G. & Charlton, C., (2013) ‘runmlwin - A Program to Run the MLwiN Multilevel Modelling Software from within Stata’, *Journal of Statistical Software,* 52 (11), pp.1-40.

Leckie, G., (2013) ‘Module 11: Three-Level Multilevel Models, Stata Practical’, *Centre for Multilevel Modelling*. Available at: <https://www.cmm.bris.ac.uk/lemma> *Accessed 15/3/2015.*

Lee, Y. & Gordon, R., (2005) ‘Tax Structure and Economic Growth’, *Journal of Public Economics*, 89(5), pp.171-190.

LeMay-Boucher, P. & Dagnelie, O., (2014) ‘The Divorced Spheres of Beninese Spouses’, *Journal of International Development,* Vol. 26(1), pp. 46-58.

Lewandowski, P., (2007) ‘PESCADF: Stata module to perform Pesaran's CADF panel unit root test in presence of cross section dependence’, *Statistical Software Components*.

Lincove, J.A., (2012) ‘The Influence of Price on School Enrolment under Uganda’s policy of free primary education’, *Economics of Education Review*, 31, pp. 799-811.

Lincove, J. A., (2015) ‘Improving Identification of Demand-Side Obstacles to Schooling: Findings from Revealed and Stated Preference Models in Two SSA Countries’, *World Development*, 66, pp. 69-83.

Martinez-Mongay, C., (2000) ‘ECFIN's effective tax rates: properties and comparisons with other tax indicators’, European Commission, Directorate-General for Economic and Financial Affairs.

McIntosh, C., (2008) ‘Estimating Treatment Effects from Spatial Policy Experiments: An application to Ugandan microfinance’, *Review of Economics and Statistics*, 90(1), pp. 15-28.

McNabb, K. & LeMay-Boucher, P., (2014) ‘Tax Structures, Economic Growth and Development’, ICTD Working Paper 22(September), Brighton: International Centre for Tax and Development.

Mendoza, E. G., Razin, A. & Tesar, L. L., (1994) ‘Effective tax rates in macroeconomics:

Cross-country estimates of tax rates on factor incomes and consumption', *Journal of Monetary Economics*, 34(3), pp. 297-323.

Mendoza, E., Milesi-Ferretti, G. & Asea, P., (1997) 'On the ineffectiveness of tax policy in altering long-run growth: Harberger's super neutrality conjecture', *Journal of Public Economics*, 66(1), pp. 99-126.

Myles, G., (2007) 'Economic growth and the role of taxation', report prepared for the OECD, *University of Exeter and Institute for Fiscal Studies.*

ODI, (2011) 'Benin's Progress in Education: Expanding access and closing the gender gap', *Overseas Development Institute.*

Paxton, J., Graham, D. & Thraen, C., (2000) 'Modeling Group Loan Repayment Behavior: New Insights from Burkina Faso', *Economic Development and Cultural Change*, 48(3), pp. 639-655.

Pesaran, M., (2006) 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica*, 74(4), pp. 967–1012.

Pesaran, H.M., Shin, Y. & Smith, R., (1999) 'Pooled Mean Group Estimation of Dynamic Heterogeneous Panels', *Journal of the American Statistical Association*, 94:446, pp. 621-634.

Prichard, W., (2016) 'Reassessing Tax and Development Research: A New Dataset, New Findings, and Lessons for Research', *World Development*, Vol. 80, pp.48-60.

Prichard, W., Goodall, A. & Cobham, A., (2014) 'The ICTD Government Revenue Dataset', ICTD Working Paper 19, Brighton: International Centre for Tax and Development.

Putnam, D.P., Leonardi, R. & Nanetti, R.Y., (1993) 'Social Capital and Institutional Success', in *Making Democracy Work: Civic Traditions in Modern Italy,* Chapter 6, Princeton, NJ: Princeton University Press, pp.163-185.

Rasbash, J., Charlton, C., Browne, W.J., Healy, M. & Cameron, B., (2015) 'MLwiN Version 2.35', Centre for Multilevel Modelling, University of Bristol.

Riezman, R. & Slemrod, J., (1987) 'Tariffs and collection costs', *Weltwirtschaftliches Archiv*, 123(3), pp. 545-549.

Rodriguez, G. & Goldman, N., (2001) 'An Assessment of Estimation Procedures for Multilevel Models with Binary response', *Journal of the Royal Statistical Society*, Series A, 158(1), pp. 73-89.

Sandefur, J. & Glassman, A., (2015) 'The political economy of bad data: evidence from African survey and administrative statistics', *The Journal of Development Studies*, 51(2), pp. 116-132.

Schneider, F., Buehn, A. & Montenegro, C., (2010) 'New Estimates for the Shadow Economies all over the World', *International Economic Journal*, 24(4), pp. 443-461.

Sharma, M. & Zeller, M., (1997) 'Repayment Performance in Group – Based Credit Programs in Bangladesh: An Empirical Analysis', *World Development*, Vol. 25(10), pp. 1731 – 1742.

Skrondal, A. & Rabe-Hesketh, S., (2004) 'Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models', Boca Raton, FL: Chapman & Hall/CRC.

Stiglitz, J. E., & Weiss, A., (1981) 'Credit rationing in markets with imperfect information', *American Economic Review*, Vol. 71(3), pp. 393-410.

Summerfield, H., (1996) 'A Note on ROSCAs among Northern Somali Women in the United Kingdom', *Money-Go-Rounds, The Importance of Rotating Savings and Credit Associations for Women*, BERG, Oxford, UK.

Tanzi, V. & Zee, H., (2000) 'Tax policy for emerging markets: developing countries', *National Tax Journal*, 53(2), pp. 299-322.

Thirlwall, A., (2006) 'Growth and Development 8th Edition', Basingstoke and New York: Palgrave Macmillan.

UIS, (2015) ‘Enrolment ratios’, *UNESCO Institute for Statistics*
Available at: <http://data.uis.unesco.org> *Accessed 10/9/2015.*

UN High Level Panel, (2013) ‘A new global partnership: Eradicate poverty and transform economies through sustainable development’, *Final Report of the UN High-Level Panel of Eminent Persons on the Post-2015 Development Agenda.*

UN, (2014) *The Millennium Goals Development Report*
Available at
<http://www.un.org/millenniumgoals/2014%20MDG%20report/MDG%202014%20English%20web.pdf> *Accessed 4/1/2015.*

UNESCO, (2012) ‘UIS Frequently Asked Questions. Education Statistics’
<http://www.uis.unesco.org/Education/Documents/FAQ-education-statistics-rev5-en.pdf> *Accessed 20/10/2015.*

UNSTATS, (2016) ‘Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators’, *United Nations Economic and Social Council*
Available at <http://unstats.un.org/unsd/statcom/47th-session/documents/2016-2-IAEG-SDGs-Rev1-E.pdf> *Accessed 12/3/2016.*

Van den Brink, R. & Chavas, J-P., (1997) ‘The microeconomics of an indigenous African institution: the rotating savings and credit association’, *Economic Development and Cultural Change*, Vol. 45, pp. 745–772.

Wenner, M., (1995) ‘Group Credit: A Means to Improve Information Transfer and Loan Repayment Preference’, *The Journal of Development Studies,* Vol. 32(2), pp. 263-281.

Woodruff, C., (2013) ‘Registering Tax for Growth: Tax and the Informal Sector in Developing Countries’, Chatham House & University of Warwick. Available at *<http://www2.warwick.ac.uk/fac/soc/economics/research/centres/cage/tax/chj854_cage_woodruff_bp_09_07_13_web.pdf> Accessed 15/1/2014.*

Wydick, B., (1999) ‘Can social cohesion be harnessed to repair market failures? Evidence from group lending in Guatemala’, *The Economic Journal*, 109(457), pp. 463-475.

Xing, J., (2011) ‘Does Tax Structure Affect Economic Growth? Empirical Evidence from OECD Countries’, Oxford University Centre for Business Taxation Working Paper No. 11/2.

Zeller, M., (1998) ‘Determinants of repayment performance in credit groups: The role of program design, intragroup risk pooling, and social cohesion’, *Economic Development and Cultural Change*, 46(3), pp.599-620.

Zellner, A., (1962) ‘An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias’, *Journal of the American Statistical Association*, 57, pp. 348-368.

Appendices

Appendix A. Survey Methodology

The survey was carried out in two of the poorest areas of the city of Cotonou, namely Vossa and Enagnon. Households were selected according to the following process: We obtained a map of Enagnon and performed a simple random selection of lots out of all those identified. In Vossa, a pseudo-random process was used by which every tenth lot according to a specific direction was picked. Starting points were selected to be equidistant from one another and so that they covered the whole district. In these two districts it is often the case that many households live on the same lot in semi-detached rooms. Enumerators selected one room per lot according to a clockwise selection, varying from lot to lot (for the first lot of the day they selected the first room clockwise, for the second one the second room clockwise and so on). Overall, only 3 households categorically refused to be surveyed and were replaced by other randomly selected households. Enumerators were asked to pass several times and at different times of the day, until contacts were established in such a way that none of the selected households were skipped. The most qualified of our enumerators also acted as a supervisor and visited many households already interviewed in order to check the accuracy of the responses. Other than that, we analysed every completed questionnaire closely. Several appointments were held with each team of enumerators and in cases of incoherence or lack of answers, we regularly sent them back on the field. Every household was compensated for their time with a donation of 1500 CFA (around USD 3).

Appendix B. Test for equivalence of means between those groups sampled in both 2004 and 2006 and those just in 2004

Variable	*2004 only*		*Both 2004 and 2006*		*Diff*
	Mean	*s.e.*	*Mean*	*s.e.*	
Membership size	28.980	2.787	33.175	3.445	-4.199
No. of cycles completed	8.083	1.539	8.639	2.215	-0.555
President	0.400	0.053	0.299	0.046	0.104
Paid	0.341	0.052	0.227	0.043	0.114*
*President*Paid*	0.094	0.032	0.062	0.025	0.032
Random	0.588	0.054	0.680	0.048	-0.092
Experienced problems pre-2004	0.118	0.035	0.206	0.041	-0.089
Written Rules	0.647	0.052	0.598	0.050	0.049
Monthly meetings	0.377	0.053	0.371	0.049	0.026
More severe sanctions on delinquent member	0.694	0.050	0.536	0.051	0.158**
Pot size (1000's of CFA)	124.681	13.452	123.288	16.919	1.394
Single ethnicity	0.212	0.045	0.237	0.043	-0.025
Only men	0.294	0.049	0.237	0.043	0.057
Only woman	0.200	0.044	0.155	0.037	0.045
Survived Past Problems	0.118	0.035	0.206	0.041	-0.089
Started with					
Friends	0.471	0.055	0.433	0.051	0.038
Family	0.059	0.026	0.082	0.028	-0.024
Members of same trade	0.188	0.043	0.144	0.036	0.044
Neighbours	0.224	0.046	0.186	0.040	0.038
Members of another group	0.047	0.023	0.072	0.026	-0.025
Other	0.012	0.012	0.072	0.026	-0.060**
President/Committee decides	0.765	0.046	0.650	0.049	0.115*
New members must be known	0.565	0.054	0.608	0.050	-0.044
Survey on new members	0.871	0.037	0.794	0.041	0.077
Other conditions?	0.835	0.041	0.856	0.036	-0.020
N = 182 (full sample)	*N=85*		*N=97*		

p<0.1; **p<0.05;*p<0.01. Note: There are only 80 observations for the variable 'No. of cycles completed' in the '2004 only' sample, due to missing data.*

Appendix C: Replication of Table 3.3 with alternative dependent variable

Depvar: Net attendance	(1)	(2)	(3)
Individual level Characteristics			
Gender	-0.0928***		
	(0.00962)		
Age	0.346***	0.352***	0.336***
	(0.00833)	(0.0108)	(0.0111)
Age²	-0.0163***	-0.0163***	-0.0161***
	(0.000442)	(0.000572)	(0.000583)
Adopted	-0.158***	-0.0325	-0.235***
	(0.0407)	(0.0229)	(0.0524)
Worked	-0.0351**	-0.0161	-0.0532***
	(0.0158)	(0.0149)	(0.0193)
Household level Characteristics			
Household Size	-0.00218**	-0.00219**	-0.00235
	(0.00103)	(0.00105)	(0.00156)
Religion			
Christian	-	-	-
Islam	-0.0692***	-0.0826***	-0.0546***
	(0.0172)	(0.0187)	(0.0187)
Other	-0.0338***	-0.0374***	-0.0314**
	(0.0113)	(0.0122)	(0.0145)
Household Wealth Level			
Poorest	-0.281***	-0.291***	-0.291***
	(0.0170)	(0.0202)	(0.0229)
Poorer	-0.193***	-0.221***	-0.181***
	(0.0139)	(0.0181)	(0.0170)
Middle	-0.114***	-0.155***	-0.0904***
	(0.0130)	(0.0184)	(0.0166)
Richer	-0.0557***	-0.0952***	-0.0354***
	(0.0115)	(0.0184)	(0.0131)
Richest	-	-	-
Household head's education level			
None	-	-	-
Primary	0.0644***	0.0564***	0.0724***
	(0.0102)	(0.0121)	(0.0139)
Secondary	0.148***	0.143***	0.158***
	(0.0113)	(0.0142)	(0.0150)
Tertiary	0.102***	0.238***	0.0610
	(0.0279)	(0.0252)	(0.0424)
School considered essential?	0.00755	0.00387	0.0128
	(0.0127)	(0.0133)	(0.0170)
Rural	-0.0180	-0.0153	-0.0229
	(0.0119)	(0.0128)	(0.0144)
Commune level Characteristics			
Distance to school	-0.0138	-0.0269**	0.00160
	(0.0107)	(0.0109)	(0.0126)
(ln) Schools per 5-14 year olds	0.145***	0.155***	0.133**
	(0.0432)	(0.0422)	(0.0519)
Observations	26,673	14,003	12,670

*Standard errors in parentheses; *p<0.1; **p<0.05;***p<0.01*

Appendix D: Estimates of Random effects for Null model, by estimation method.

Intercept variance estimate	*Estimation method used*				
	MQL1	MQL2	PQL1	PQL2	MLE
Commune	0.323	0.326	0.390	0.573	0.688
Household	0.387	0.399	0.701	1.223	2.020

Appendix E. List of Countries Included

Low Income (GNI per capita < $1,045)

Burundi, Benin, Central African Republic, Gambia, Cambodia, Mali, Mozambique, Malawi, Niger, Nepal, Rwanda, Sierra Leone, Tanzania, Uganda, Zimbabwe.

Lower-Middle Income ($1,045 < GNI per capita < $4,125)

Bangladesh, Egypt, Ghana, Guatemala, India, Kenya, Sri Lanka, Lesotho, Morocco, Nicaragua, Philippines, Sudan, Senegal, El Salvador, Swaziland, Ukraine.

Upper-Middle Income ($4,125 < GNI per capita < $12,736)

Albania, Bulgaria, Belize, Brazil, Botswana, Congo, Republic, Costa Rica, Cuba, Dominican Republic, Algeria, Ecuador, Gabon, Iran, Jordan, Mexico, Mongolia, Mauritius, Namibia, Pakistan, Panama, Peru, Paraguay, Thailand, Tonga, Tunisia, Turkey, South Africa.

High Income (GNI per capita > $12,736)

Argentina, Australia, Austria, Belgium, Bahrain, Barbados, Canada, Chile, Chile, Cyprus, Czech Republic, Germany, Denmark, Switzerland, Spain, Finland, France, United Kingdom, Greece, Hong Kong, Hungary, Ireland, Iceland, Israel, Italy, Japan, Korea, Lithuania, Luxembourg, Malta, Netherlands, Norway, New Zealand, Poland, Portugal, Singapore, Slovakia, Slovenia, Sweden, Trinidad and Tobago, Uruguay, United States, Venezuela.

Appendix F. Summary of Disaggregated Tax shares.

Total Taxes (including Social Contributions) =

- Income Tax
  - Personal Income Tax (PIT)
    - *NB Includes Taxes on Payroll and Workforce*
  - Social Contributions (SC)
  - Corporate Income Tax (CIT)
- Consumption and Property Taxes
  - Goods and Services
    - *NB Includes 'Other' Tax*
  - Trade
  - Property

Where

(i) Income Tax + Consumption and Property Taxes = Total Taxes
(ii) PIT + SC + CIT = Income Tax
(iii) Goods and Services + Trade + Property = Consumption and Property

Appendix G. PMG Estimation, full sample with alternative time controls

Dependent Variable: Δ(log) GDP per capita.

	(1)	(2)	(3)	(4)	(5)
Physical Capital	-0.003	0.007	0.001	0.004***	0.003
	(0.013)	(0.014)	(0.014)	(0.013)	(0.014)
Human Capital	-0.017**	-0.022**	-0.020**	-0.009	-0.028***
	(0.008)	(0.010)	(0.010)	(0.008)	(0.010)
Population Growth	-0.328***	-0.535***	-0.288***	-0.457***	-0.387***
	(0.097)	(0.120)	(0.103)	(0.093)	(0.107)
Tax / GDP	0.030	0.031***	0.117***	0.068***	0.119***
	(0.022)	(0.006)	(0.025)	(0.022)	(0.027
Tax Structure Variables					
Income taxes	-0.023			-0.025*	-0.039**
	(0.014)			(0.014)	(0.017)
PIT		-0.064***			
		(0.021)			
Social Contributions		-0.077***			
		(0.025)			
CIT		0.012			
		(0.012)			
Consumption & Property					
Goods & Services			0.074***		0.020*
			(0.016)		(0.012)
Trade			0.042***		
			(0.016)		
Property			-0.027	-0.046	-0.007
			(0.055)	(0.055)	(0.055)
Omitted Tax Variable	**Consumption & Property**		**Income**	**Consumption**	**Trade**
Observations	2664	1811	2562	2562	2290
Groups	100	68	96	96	84
Time Controls					
Country Specific Time Trends	Yes	Yes	Yes	Yes	Yes
5-Year Dummies	No	No	No	No	No

*Standard errors in parentheses; *P<0.1; **P<0.05; ***P<0.01*